Creating and functionalising a synthetic compartment in plants



András Sándor

New College

Department of Plant Sciences

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy in Synthetic Biology

Hilary term 2022



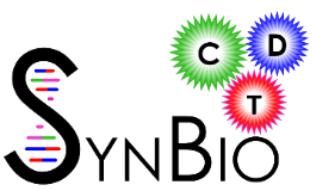


Table of contents

Table of contents 2
Table of figures 5
Table of tables 7
Acknowledgements 8
Abstract 10
Abbreviations used 11
Chapter 1: Introduction 13
1.1 Industrial biosynthesis and recombinant protein production in plants 13
1.2 Compartmentalisation of metabolic pathways 16
1.3 Microdomains are compartments utilising metabolic channelling 21
1.4 Partial remodelling of the ER for effective compartmentalisation: organised smooth endoplasmic reticulum (OSER) in plants 24
1.5 Project rationale and aims 31
Chapter 2: Materials & Methods 33
2.1 Genetic construct design 33
2.2 Molecular cloning 34
2.3 Plant growth conditions 36
2.4 Plant transient transformation 37
2.5 Phenotypic analysis of stably transformed *A. thaliana* plants 38
2.6 Purification of transiently expressed proteins from mature *N. benthamiana* leaves 39
2.7 Fluorescent protein quantification 40
2.8 Analysis of purified proteins by sodium-dodecyl-sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) 41
2.9 Western Blotting 42
2.10 Confocal laser scanning microscopy 43
2.11 Quantitative analysis of ER network architecture and dynamics 44
2.12 Transmission electron microscopy (TEM) 45
2.13 Gas chromatography-mass-spectrometry (GC-MS) metabolite identification and quantification 45
Chapter 3: Characterisation of a synthetic, self-assembling de novo compartment in plant cells 54
3.1 Introduction 54
3.2 Results 56
3.2.1 Design of modular compartment forming protein variants 56

3.2.2 Initial characterisation of the six compartment forming constructs 57
3.2.3 High-resolution analysis of compartment structures 65
3.2.4 Compartment topology and diffusion dynamics 68
3.2.5 Quantitative ER network architecture and dynamics 74
3.3 Discussion 77
3.3.1 The effect of the orientation of the dimerization domains on the morphology of the compartments 77
3.3.2 Comparison between OSER and the synthetic compartments 80
3.3.3 Potential applications of the synthetic compartment 83
Chapter 4: Further investigation into the relationship between the structure of the synthetic compartments and the constitutive domains of the compartment-forming protein constructs ... 85
4.1 Introduction 85
4.2 Results 87
4.2.1 Design of the second-generation compartment forming protein constructs 87
4.2.2 Initial characterisation of the second-generation constructs 91
4.2.3 High-resolution analysis of second-generation compartments 97
4.2.4 Compartment diffusion dynamics 100
4.2.5 ER network dynamics 101
4.3 Discussion 106
4.3.1 The viability of the second-generation synthetic compartments for potential applications 106
4.3.2 The hypothesis of an ER - peroxisome - vacuole pathway for constructs not forming a compartment 107
4.3.3 The effect of the protein construct's modular domains on the morphology of the compartment 109
Chapter 5: Foundation steps towards application 117
5.1 Introduction 117
5.2 Results 119
5.2.1 Developing a method to purify the compartment-forming protein constructs 119
5.2.2 Proof-of-principle purification of a genetically fused, compartment-localised enzyme 122
5.2.3 Assessing the efficiency of the compartment as a custom surface area for anchoring proteins 124
5.2.4 Purification of enzymes anchored to the surface of the compartment 130
5.2.5 Testing the functionality of enzymes attached to the surface of the compartment ... 132
5.2.6 Characterisation of stably transformed *A. thaliana* lines 137
5.3 Discussion 144

5.3.1 The viability of a synthetic ER-derived compartment as a platform for improving plant-based recombinant protein production and as a metabolic microdomain 144
5.3.2 The compartment's potential protective capacities for attached proteins from proteases 148
5.3.3 Phenotypical differences of stably transformed *A. thaliana* lines and their consequences for application 150
Chapter 6: Discussion 152
6.1 Short summary of results and future perspectives on the compartment system 152
6.2 Potential further investigations into the basic properties of the ER-derived synthetic compartment system 154
6.2.1 Incorporating the advancements in *in silico* modelling of proteins and protein-lipid interactions into the design pipeline of compartment-forming constructs 154
6.2.2 The use of novel construct domains and techniques to investigate compartment formation and structure 156
6.2.3 The possibility of exploring other hosts for the compartment system 159
6.3 Exploring the established and potential further applications of the compartment system 161
6.3.1 Production of difficult-to-express proteins 161
6.3.2 Building a platform for a synthetic metabolic microdomain 163
6.3.3 Transitioning to a stably expressing host to improve all putative applications 166
6.3.4 Potential avenues for further long-term improvements to the compartment system for novel applications 167
6.4 Comparison of the compartment system to the currently available methods of compartmentalisation 171
References 173
Appendix A: Supporting Information 196

Table of figures

Figure 1.1: The benefits of probabilistic metabolic channelling in a microdomain. 22

Figure 1.2: Examples of OSER morphologies in plant cells. 25

Figure 1.3: A unified theory of OSER formation. 28

Figure 1.4: Remodelling the ER using the synthetic oligomerising G22Y polyprotein. 30

Figure 3.1: Schematic of the compartment forming polyprotein variants used. 57

Figure 3.2: Confocal microscopy of the G22Y family constructs 5 days after transformation of mature tobacco leaves. 60

Figure 3.3: Co-expression G22C and G22 with fluorescent subcellular markers in mature tobacco leaves. 61

Figure 3.4: Co-expression of ER markers with the synthetic compartments, to determine if the compartments inherit their lumen and membrane from the ER. 63

Figure 3.5: Expression time-course of G22Y from 24 h to 6 weeks in mature tobacco leaves. 63

Figure 3.6: High-resolution airy-scanning confocal microscopy images of the G22Y family of constructs 5 days after transformation of mature tobacco leaves. 66

Figure 3.7: TEM of 22Y and C22Y expressing tobacco epidermis cells. 67

Figure 3.8: A representative example of FRAP of both the lumenal ER marker RFP-HDEL within and the YFP signal of the synthetic compartment. 69

Figure 3.9: Retarded lumenal and cytosolic recovery rate in compartments after FRAP. 70

Figure 3.10: Visualising the cytosol trapped within the membranes of the compartments. 71

Figure 3.11: Co-expression of tonoplast and ER markers with 22Y spheres (yellow) in mature tobacco leaves. 73

Figure 3.12: Principal component analysis of the ER network dynamics. 75

Figure 3.13: Relative C-terminal positioning of dimerising proteins affects the structure of the synthetic compartment. 79

Figure 4.1: Schematic of the second-generation compartment forming polyprotein constructs investigated. 89

Figure 4.2: Confocal microscopy of the second-generation constructs 5-7 days after transformation of mature tobacco leaves. 94

Figure 4.3: Co-expression of ER markers with the compartments to determine if the lumen and membrane of the compartments are derived from the ER. 96

Figure 4.4: High-resolution airy-scanning confocal microscopy images of second-generation compartment forming constructs 5-7 days after transformation of mature tobacco leaves. 99

Figure 4.5: Retarded cytosolic fluorescence recovery rate of the small molecule probe FDA in compartments after FRAP. 101

Figure 4.6: Principal component analysis of the effects of C22Y-SpyT-his on ER network dynamics. 102

Figure 4.7: Principal component analysis of the effects of the tetrameric constructs on ER network dynamics. 104
Figure 4.8: The general hypothesis on the driving forces and checkpoints for the ER-derived compartment formation. 116
Figure 5.1: Purification of the compartment-forming scaffold proteins. 121
Figure 5.2: Purification of a genetically fused enzyme via immunoprecipitation. 123
Figure 5.3: Schematic representation of SpyCatcher – SpyTag binding to the surface of the compartment. 127
Figure 5.4: Visualisation of proteins anchored to the cytosolic surface of the compartment via SpyCatcher – SpyTag covalent binding. 128
Figure 5.5: Efficient recruitment of proteins to the cytosolic surface of the compartment via SpyCatcher – SpyTag covalent binding. 129
Figure 5.6: Purification of a covalently bound enzyme – protein-forming scaffold complex. 131
Figure 5.7: Testing the efficiency of a metabolic microdomain via the 2,2'-biphenol breakdown pathway. 133
Figure 5.8: Visualisation of anchoring an enzyme to the surface of the compartment. 136
Figure 5.9: The activity of an enzyme anchored to the surface of the compartment. 136
Figure 5.10: Fluorescent whole-seedling imaging of stably transformed A. thaliana. 138
Figure 5.11: Confocal microscopy of T2 stably transformed A. thaliana lines. 139
Figure 5.12: High-resolution airy-scanning confocal microscopy of T2 stably transformed A. thaliana lines. 140
Figure 5.13: Stably transformed T2 A. thaliana phenotypes. 142
Figure 5.14: Graphical representation of the statistical analysis of the key macro-phenotypes. . 143
Figure 6.1: Potential improvements to the compartment system to enable further applications. 169
Supplementary Figure S1: The citramalate – ribitol standard curve. 196
Supplementary Figure S2: Stable 2-week-old compartments. 197
Supplementary Figure S3: Transiently transformed leaves show no negative macroscopic phenotypes. 197
Supplementary Figure S4: Helical structures observed in G22Y compartments. 198
Supplementary Video SV1: Z-stack of 22Y structures. 199

Table of tables

Table 2.1: Gene parts used for the design of the genetic constructs. 49

Table 2.2: PCR and sequencing primers used to generate constructs. 53

Table 3.1: Parameters used in the AnalyzER software package to investigate network dynamics. 75

Table 4.1: The parameters contributing the greatest proportion of variance to PC1 during the AnalyzER analysis of tetrameric constructs. 105

Table 4.2: Categorisation of compartment morphologies and constructs. 112

Table 5.1: Results of statistical analyses of key macro-scale phenotypes. 142

Supplementary Table S1: Individual ANOVA of every parameter used in AnalyzER testing in Chapter 3.2.5.. 200

Acknowledgements

Finishing a PhD is a long journey, so there is a substantial list of people to thank for their help. First and foremost, I want to thank Lee Sweetlove for his guidance and supervision over the last five years. Likewise, I would like to thank Mark Fricker for his advice and help throughout the project. I am especially grateful that my two supervisors gave me free reign to explore the exciting parts of my project, making it a thoroughly enjoyable experience to work with them. I really think we produced some very exciting research, and I will always remember my time at Oxford fondly.

A number of people also gave very valuable input on technical aspects and future perspectives of the project, so I am grateful to Verena Kriechbaumer, Chris Hawes, Philippe Varennes-Jutras, Nick Kruger and Markus Schwarzlander. Lee, Mark and Verena deserve my extra thanks for editing my literature review on plant OSERs, which formed the basis of a small part of this thesis's introduction too.

I am thankful to everyone who provided help with various methods: our beloved lab manager Pedro Bota for his help with dozens of methods, especially with the GC-MS experiments, Clare Jones for her assistance with preparing samples for TEM, Charlotte Pain for advice on using AnalyzER and Errin Johnson who carried out TEM sectioning and imaging as a service. Special thanks to Anya Lindström-Battle for her work on tetrameric coiled-coil structures, some of which ended up in this thesis, which I was lucky enough to supervise. I am also grateful for Naomi Cannell for her suggestion of the biphenol breakdown pathway as the perfect candidate for testing a microdomain. A number of people were generous enough to provide probes, strains or other tools for my project, including Mark Howarth who provided the SpyCatcher system, Philippe Varennes-Jutras who gifted me the AGL1 *Agrobacterium* lines used in this project, Verena Kriechbaumer and Beatriz Moreno García who provided numerous fluorescent probes and Federica Brandizzi who collaborated with us to generate the stably transformed *Arabidopsis* lines. I am also thankful to Ian Moore and Marketa Samalova for their preliminary experiments on the G22Y compartments that made this project possible in the first place. I am also grateful for the funding provided by the Synthetic Biology CDT and EPSRC, which made my research possible.

Of course, I have to also thank the members of the Sweetlove group, Sanu Shameer, Lottie Chapman, Mann Konuntakiet, Nadine Töpfer, Beatriz Moreno García, Ed Smith, Naomi Cannell, Alex Bones, Corinna Hartinger, Anya Lindström-Battle and Hilary Hunt, for their help throughout the years, their camaraderie and generally providing an awesome environment to work in!

I would like to give special thanks to my family and friends for keeping me sane and for their unwavering support these past years – listing all of you would take up another whole page, so I will spare the reader that, but know that I am grateful!

Finally, I wish to acknowledge the exactly 578 plants (yes, I counted) that gave their lives during the 4 years of this project. Since in science, the loss of plant lives is usually much less appreciated than of animal lives, I am dedicating this thesis to them. Hopefully your sacrifice has not been in vain.

Abstract

Compartmentalisation is a universal feature of life, allowing organisms to separate their functions from each other and the environment. It enables greater control and complexity for cellular processes. Taking a cue from nature, engineered compartmentalisation has become one of the top priorities of synthetic biology. While many alternative strategies have emerged, building a synthetic membranous organelle *de novo* remains prohibitively difficult, despite its potential as a tool for metabolic engineering. This lack of an orthogonal synthetic organelle especially hampers the development of the nascent plant molecular farming industry. In this thesis, I describe the design and characterisation of a novel synthetic membranous compartment derived from the plant endoplasmic reticulum (ER), and proof-of-principle experiments to showcase its potential as a bioproduction platform in plants. The formation of the compartment is triggered by overexpression of a single recombinant scaffold protein, and I investigated the relationship between the structure of the scaffold protein and the morphology of the resultant compartment. I also compared the compartment to a conceptually similar ER rearrangement called organised smooth endoplasmic reticulum to determine their relation. Notably, the formation of the compartment had very little impact on the overall health of the host plant, and successful pilot experiments for the purification of the scaffold and other proteins anchored to it, as well as running enzymatic reactions on the surface of the compartment demonstrate the great potential of this system to serve as a plant synthetic biology tool. With further development, this platform could be a valuable asset to the plant biosynthesis and recombinant protein production industry.

Abbreviations used

Abbreviation	Full form
ER	endoplasmic reticulum
OSER	organised smooth endoplasmic reticulum
ESCRT	endosomal sorting complexes required for transport
GFP	green fluorescent protein
HMGR	3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase
YFP	yellow fluorescent protein
mRFP	monomeric red fluorescent protein
PCR	polymerase chain reaction
LB	Lysogeny Broth
SDS-PAGE	sodium-dodecyl-sulphate-polyacrylamide gel electrophoresis
APS	ammonium persulphate solution
TEMED	N, N, N', N' Tetramethylethylene-1,2-diamine
HRP	horseradish peroxidase
FDA	fluorescein diacetate
FRAP	fluorescence recovery after photobleaching
PCA	principal component analysis
PC	principal component
MANOVA	multivariate analysis of variance
ANOVA	analysis of variance

TEM	transmission electron microscopy
GC-MS	gas chromatography-mass-spectrometry
WT	wild type
FDH	formate dehydrogenase
Luc	luciferase
HbpA	2-hydroxybiphenyl-3-monooxygenase
HbpC	extradiol ring cleavage dioxygenase
MdCMS	citramalate synthase
ERAD	ER-associated degradation
MCS	membrane contact sites
FRET	Förster resonance energy transfer
TEV	Tobacco Etch Virus
F2A	foot-and-mouth disease virus 2A peptide

Chapter 1: Introduction

1.1 Industrial biosynthesis and recombinant protein production in plants

The use of plants as an industrial platform for the production of high value biochemicals (referred to as molecular farming[1]) has been gaining steady traction since the first proof-of-principle studies showcasing recombinant antibodies generated in plants in the late 1980s[2].

Forty years later, there are more than 30 companies worldwide involved in molecular farming, from small start-ups to major industrial players[3]. Antibodies remain a common target, with an HIV-specific monoclonal antibody becoming the first plant-based antibody to pass phase I clinical trials[4]. Molecular farming is also used for producing other pharmaceuticals, such as cheap edible vaccines trialled for both animals[5] and humans[6], and recombinant plasma proteins (e.g., complement factor H[7]) and enzymes (e.g., taliglucerase alfa for Gaucher's disease[8]) for replacement therapy[9,10]. Enzymes produced in plants are also used in various industries: trypsin for food processing and leather tanning, α-amylase for bioethanol production and spider-silk protein as an industrial material[11,12] are just some of the commercially available examples. Moreover, biological research frequently uses plant-produced reagents, like albumin, lysozyme, collagen and various growth factors[12].

Beyond proteins, industrial biosynthesis in plants is also focused on the large-scale production of high value metabolites. Humans have exploited the various specialised compounds produced by plants for their medicinal properties for thousands of years, but most of these valuable secondary metabolites are produced at low levels naturally[13]. Research into the production of these compounds usually focuses on transferring entire

pathways from non-model organisms into hosts more suitable for molecular farming[14]. Similar to proteins, expensive secondary metabolites with medicinal properties are of the most interest for industrial production. The antimalarial drug artemisinin[15] and the anticancer drug taxol[16,17] are some of the most known examples to be reconstituted in novel plant hosts and commercialised. More uncommon uses of plant metabolic engineering include the marketing of a moss species as a house plant that is genetically modified to produce terpenoids used in fragrances (e.g: patchoulol and linalool) for a pleasant smell[7]. Research into the improved production of secondary metabolites for medicinal and industrial uses is continuous, with more plant-produced pharmaceuticals and industrial compounds becoming available every year. Further targets include food colourings, cosmetics and medicine from various groups of metabolites (such as carotenoids, sterols, anthocyanins, alkaloids[14,18]) and even some primary metabolites (such as succinate, fumarate or even butanol[19]) that can serve as precursors for the synthesis of other high-value biochemicals or biofuels, replacing fossil-fuel derived compounds[20].

The growing interest in molecular farming is no accident: plants have a number of notable benefits over other production platforms. Modern agricultural methods make large scale plant cultivation for bioproduction in greenhouses significantly less costly and more scalable than microbial fermentation-based systems. Open-field growth has an even greater competitive advantage in terms of scalability and costs[1,21], but the regulation of genetically modified plants, especially for open-field growth is a controversial issue, hindering commercial uptake[9,20]. Another advantage is that plants are autotrophic, reducing the cost of nutrient resources required and the ecological footprint of their growth when compared to microbial fermentation or mammalian cell lines[14]. Plants are also eukaryotic with a similar secretory pathway to mammalian cells and the innate

capacity for the correct folding of proteins with complex secondary and tertiary structures and post-translational modifications (e.g., glycosylation[12,21]). Compared to animal cell cultures, plants have a lower risk of infection or contamination, and the products are animal-free which also carries a minor but increasingly important marketing benefit[1,9,22]. Furthermore, many important secondary metabolites of commercial interest are produced natively in plants, so reconstituting these pathways in other plant species for improved production is generally easier than doing so in hosts from different kingdoms[14,18].

Although plant synthetic biology is much less standardised than bacterial, yeast or mammalian alternatives, a number of well-characterised model crops are readily available (e.g., *Arabidopsis thaliana* or *Nicotiana tabacum)*, and hundreds of different high value biochemicals have already been successfully produced in these[1]. Whole-plant expression of target proteins or metabolites currently dominates the field, but plant cell cultures are also becoming competitive due to significant advances recently, such as plant cell packs for high-throughput screening[23,24]. Cell-free tobacco BY-2 lysate-based systems could also gain momentum in the future, owing to their high production capacity[25].

While the development of stable plant lines can be more time-consuming than for other established model systems due to longer generation times and more complex transformation protocols, an alternative is available for rapid recombinant protein production: *Agrobacterium tumefaciens* mediated transient transformation of mature leaves which can generate harvestable proteins within days[18,21]. Transient expression has been applied to industrial-scale molecular farming, and it is now feasible to scale up production to thousands of kilograms of tobacco leaves expressing the desired protein at high yields[26,27]. This speed can be extremely beneficial in emergency situations. For

example, plants were used to rapidly produce antibodies to treat Ebola patients during the 2014 West Africa outbreak, and a similar use was investigated in the early months of the current Covid-19 pandemic[28]. There is now a nascent industry focusing specifically on this subfield of molecular farming, such as Leaf Expression Systems in Norwich, UK[3].

However, the spread of commercialised molecular farming has been slowed by some limiting factors. Previously, there was a lack of precise genome engineering tools, although there have been recent advances in this area[14,24,29]. Furthermore, a notable issue with plants affecting the yield of target compounds, is that the desired product can accumulate inside the vacuole which can in some cases cause the breakdown of the products[30]. Although the ease of the upscaling of production can counter product loss to a certain degree, other methods for alleviating this problem are desired. One such solution to reduce product loss could be to redirect the target products or entire pathways into a more easily manipulatable subcellular compartment[31]. Indeed, a modular, easy-to-use platform for compartmentalisation in plant cells could be one of the possible next breakthroughs that could rapidly expand industrial biosynthesis and recombinant protein production in plants.

1.2 Compartmentalisation of metabolic pathways

Subcellular compartments are a fundamental principle of all life on Earth, encompassing a broad range of structures which provide a way of spatially separating matter and functions from the environment and other compartments[32,33]. Intracellular compartments play a role in virtually all aspects of life, from unicellular organisms to the most complex multicellular organisms. Even prokaryotes which were historically thought to have very little internal compartmentalisation, have been shown to use membraneless analogues of organelles for many processes, such as the nucleoid aiding DNA replication and cell division[34], or protein-

shell-based bacterial microcompartments enabling the catabolism of toxic compounds[35,36]. Eukaryotes also use membraneless organelles[37–40], such as the nucleolus for ribosome biogenesis. However, eukaryotes are most known for their use of internal membrane structures to separate cellular functions into distinct compartments, specialising for metabolism (e.g., mitochondria and chloroplasts), information storage (e.g., nucleus), protein production and breakdown (e.g., endoplasmic reticulum (ER) and peroxisomes) and many other functions[32].

The ubiquitous nature of compartmentalisation in life can be explained by its benefits. Compartmentalisation offers a way for the cells to exert spatiotemporal control over their resources and functions (e.g., by modulating substrate availability)[40]. It allows incompatible metabolic pathways (whether for mutually exclusive solution requirements, competition for substrates, or the danger of futile cycling) to exist in parallel within the cell by physically separating and quasi-independently regulating them[41]. Finally, many metabolic processes generate toxic intermediates or final products which must be sequestered from the bulk of the cell as a matter of self-protection[42].

Because of the benefits of compartmentalisation, the creation of engineered compartments has been a goal of synthetic biology since the inception of the field[31,35]. In general, three main strategies have emerged: (i) the modification of existing compartments usually by re-targeting enzymatic pathways to them; (ii) building membraneless organelles most commonly with liquid-liquid phase separation of proteins; and (iii) development of novel synthetic membranous compartments.

The utilisation of native compartments is generally considered to be the simplest approach, as enzymes can be often re-targeted with the addition or removal of signal peptides.

However, some unintended problems can arise. For example, localisation tags are not always fully reliable leading to mis-targeting of introduced proteins[43]. Furthermore, introduction of new metabolic pathways into compartments can disrupt native functions and the introduced pathways have to compete with pathways optimised by millions of years of evolution[31,44]. Nevertheless, this strategy has produced many successes in recent years, and many different organelles have been explored as new hosts for metabolic pathways, such as peroxisomes[44–47], ER and Golgi[48], vacuole[49], cell wall[50] and mitochondria[51,52]. This approach has produced notable results in plants as well, using a variety of organelles. For example, in *Nicotiana benthamiana,* terpene production has been repeatedly improved by retargeting the pathway enzymes to mitochondria[53–55], in one case producing a fifteenfold production increase[53]. Dhurrin[56] and artemisinin[57] production were increased significantly when moved to the thylakoid membranes of the chloroplast, and moth sex pheromones produced in the ER were successfully used as an *in vivo* pesticide[58].

An alternative method of compartmentalisation is to build membraneless organelles. This strategy is based on the principle that many proteins with intrinsically disordered regions will undergo liquid-liquid phase separation (demixing) and form a liquid droplet that functions as a membraneless organelle[59,60]. Beyond intrinsically disordered regions, oligomerising domains and nucleic acid scaffolds are also capable of inducing demixing[39,61]. The formation and dissolution of membraneless organelles is highly dynamic, depending on the cellular microenvironment (e.g., temperature, pH, salinity, etc.)[59,60] and often a single overexpressed protein is sufficient to trigger their formation[61,62]. The lack of a delimiting membrane means that controlling permeability does not require the inclusion of specific channels or pores. Conversely it also means that permeability cannot be controlled, and local environment of the compartment remains similar to the bulk aqueous

solution[31,63]. This lack of control of permeability necessitated the development of alternative control strategies, such as inducible formation of the compartment using optogenetic methods[61,64], or cargo proteins with inducible capture to and release from the membraneless organelle[65,66]. Despite the relatively recent increase in interest in membraneless organelles, they have already been used for metabolic engineering (e.g., with an elegant experiment to control fluxes in the deoxyviolacein synthesis pathway[64]), to control cellular processes (e.g., proliferation of yeast cells[66]) and even to introduce novel functions into the host cell (e.g., the site-specific recruitment of ribosomes and tRNA to enable orthogonal translation with non-canonical amino acids[67]) showcasing their great potential. Recently, repurposing of the native membraneless organelle called the pyrenoid (localised in the chloroplasts of algae and some liverworts for concentrating CO_2[68]) has been suggested as a strategy with significant future potential[69].

Beyond liquid-liquid phase separation, membraneless organelles can also be built by encapsulating functions inside a compartment with a proteinaceous shell. Virus-like particles can be used for this purpose, with genetically fused enzymes inside the capsids previously shown to function[70,71], but these are limited in size (and thus the number of enzymes that can be loaded inside them) and can be harmful to the hosts[72]. Bacterial microcompartments reconstituted in heterologous hosts can also serve as membraneless organelles, with larger (up to 500 nm[31]) selectively permeable shells that can be readily targeted to with the use of native bacterial microcompartment localisation tags[73]. However, these usually require a number of different proteins to form, making them more complex to utilise when compared to liquid-liquid phase separation-based compartments[31]. Nevertheless, there are examples of bacterial microcompartments being used to encapsulate metabolic activity in yeasts[74] and plants[75,76]. In tobacco chloroplasts,

CO_2 concentrating carboxysomes (a bacterial microcompartment from photosynthetic bacteria) containing ribulose bisphosphate carboxylase-oxygenase (Rubisco) and carbonic anhydrase was successfully reconstituted and shown to be functional after the endogenous Rubisco large subunit was knocked out and plants remained viable and capable of assimilating CO_2 and growth[76].

Developing a novel synthetic membranous organelle is considered to be the most challenging compartmentalisation strategy. Due to their synthetic nature, these compartments are expected to have less crosstalk with the host background, giving more freedom to metabolic engineers. The presence of a delimiting membrane would alleviate most of the downsides of a membraneless organelle, such as lack of control over the contents of the compartment. However, such a strategy requires a more thorough understanding of the biology of the host and exploiting native functions for a novel purpose can be difficult[31]. For example, the membrane biogenesis pathway is not yet fully understood and hijacking it to *de novo* generate membranous structures for specific functions is challenging[77]. Even when a novel membranous compartment is successfully generated, it requires the introduction of specific channels, receptors or other targeting systems to enable trafficking into the new compartment[77]. Despite the inherent difficulty of this approach, there is progress being made. In *Escherichia coli* a new artificial membrane vesicle trafficking system was developed as a single-step concentration and purification method for ß-carotene via a systematic approach to modulate the bacterial membrane synthesis pathway[78]. In yeast, ER-derived membranous vesicles were generated by overexpressing a scaffold protein partially derived from the maize storage protein γ-Zein, and after genetically fusing enzymes to the scaffold, the *cis,cis*-muconic acid production pathway was reconstituted inside the vesicles[77]. Flaviviruses, which naturally

produce ER-derived vesicles might also potentially be exploited to generate compartments with novel functions[79]. A related area of interest is the creation of artificial membrane-bound proto-cells with spatially segregated multi-enzyme cascades have been created[80], and the available suite of tools for increasing the complexity of these systems is steadily increasing[81].

1.3 Microdomains are compartments utilising metabolic channelling

A special case of compartmentalisation is the use of scaffolds to build microdomains, concentrating a large number of enzymes from a single metabolic pathway. Microdomains potentially improve the efficiency of metabolic pathways via probabilistic metabolic channelling, a process which facilitates the transfer of an intermediate to the subsequent enzyme in the pathway, reducing the possibility of it first diffusing to the bulk aqueous solution[69,82,83]. Metabolic channelling is widespread in nature in both the central metabolism[84–87] and the production of secondary metabolites[88–91].

Unlike direct metabolic channelling (e.g., intramolecular tunnels[92–94], electrostatic bridges[84,85] and swing arms[95,96]) which relies on direct transport of an intermediate between enzyme pairs, microdomains utilise probabilistic metabolic channelling[83] (Figure 1.1), often dynamically forming in response to external or internal cues[97]. When formed, microdomains concentrate large amounts of enzymes catalysing subsequent reactions of a metabolic pathway, increasing the probability of an intermediate encountering the next enzyme of the pathway, instead of equilibrating to the bulk aqueous solution[98]. This probability is directly influenced by the size of the microdomain, but not by the internal organisation of the enzymes[98,99].

Metabolic channelling has evolved multiple times due to its numerous benefits (Figure 1.1). As intermediates are rapidly used by subsequent steps of the metabolic pathway, they are less likely to cause negative feedback loops[100], and are sequestered from the bulk cellular environment, which can be useful when the intermediates are unstable, toxic, or secreted from the cell[83,91]. Metabolic branch points can be bypassed, increasing flux through the channelled pathway, shielding it from competing side-reactions[83,91]. Metabolic channelling also reduces the time required for the pathway to reach steady state (as the local substrate concentration is increased due to the reduced reaction volume), but it does not increase the reaction rate itself at steady state, unless the reaction is diffusion limited, which is considered to be rare[83,101].

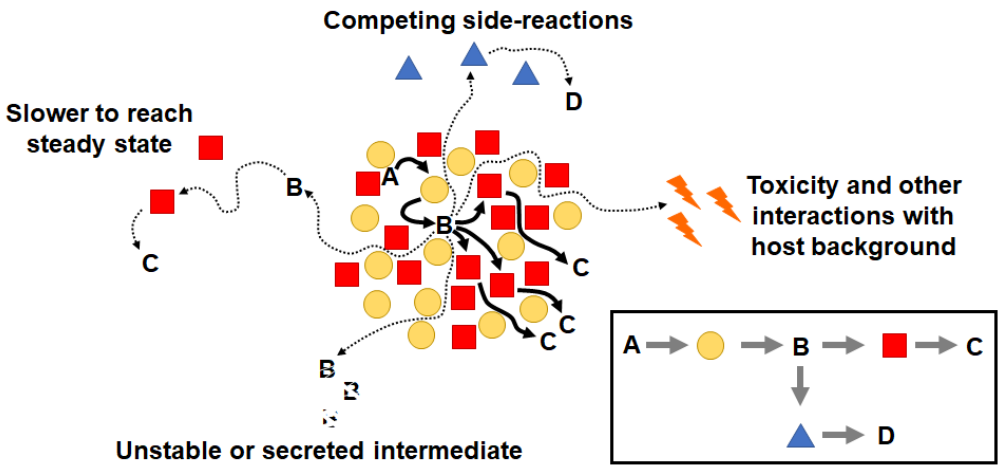


Figure 1.1: The benefits of probabilistic metabolic channelling in a microdomain. When a large number of enzymes from a subsequent pathway (a hypothetical pathway schematic is show in the box) are concentrated, the product ('B') of the first enzyme (yellow circle) has a greater probability of encountering the subsequent enzyme in the pathway (red square) and being turned into the final product 'C' than diffusing to the bulk of the cell. This probabilistic metabolic channelling means that the flux is increased towards that pathway as the intermediate B is less likely to encounter enzymes facilitating competing side-reactions (blue triangle) and become a different product 'D', and the channelled reaction is quicker to reach steady state. The rapid use of the intermediate also reduces the risk of secretion from the cell, breakdown (if the intermediate is unstable) or toxicity to the cell. Note that direct metabolic channelling (not depicted) carries similar benefits.

Engineering of direct metabolic channelling requires in-depth understanding of substrate-enzyme interactions and the molecular structure of the enzymes involved, making it difficult, although not impossible; synthetic swing arms[102] and engineered electrostatic guidance[103] have been successfully constructed. However, simply fusing pairs of enzymes is usually not sufficient to elicit direct metabolic channelling[69], because it does not facilitate direct transfer of a metabolite from one active site to the other. As such, building synthetic microdomains using multimeric scaffolds has emerged as a popular strategy to introduce probabilistic metabolic channelling to pathways of interest, since it requires less precise protein engineering while retaining the benefits of metabolic channelling[69,82].

However, membranous compartments still hold several benefits over microdomains. While intermediates are less likely to diffuse out of a microdomain, the final product will equilibrate with the bulk cellular environment if there is no membrane to retain it[31]. Membranous compartments are capable of shielding a large number of metabolic pathways from the rest of the cell, while in microdomains a high enzyme concentration is essential, so introducing multiple pathways (thus diluting all pathway enzymes) reduces the effects of metabolic channelling[83]. It is important to note though, that natural microdomains often make use of membranes as assembly surfaces[104]. Taking this cue from nature, many synthetic microdomains also utilise multimerising scaffolds that are attached to membranes[69,82,105], for example to the surface of the ER. As such, there is considerable interest in the engineering of the ER to improve its suitability to serve as a binding surface for synthetic microdomains.

1.4 Partial remodelling of the ER for effective compartmentalisation: organised smooth endoplasmic reticulum (OSER) in plants

The ER is a highly dynamic organelle with the largest membrane surface area in the cell, which undergoes significant morphological changes over time in response to endogenous and exogenous cues[106,107]. Local rearrangements of tubules and cisternae occur continuously on the scale of seconds, and large-scale morphological restructuring (e.g., during development the proportion of cisternae and tubules changes) over hours or days is also common[108,109].

Mechanistically, the architecture of the ER is influenced by ER-membrane resident proteins. One class of these proteins are the reticulons, which, with their wedge-shaped transmembrane domains, disrupt the shape of the surrounding membrane layer, thus inducing a local curvature and promoting the formation of tubules[110–112]. Lunapark proteins have an opposite effect and induce the formation of cisternae in plant ER[113]. Another family of ER shaping proteins, the atlastins, are involved in membrane fusion and the formation of three-way junctions in the ER network[114,115]. These classes of ER-shaping proteins provide the ER with significant innate plasticity. This allows cells to tolerate considerable change in ER structure, and for specialised cells to form distinct ER morphologies: for example, cells expressing a high load of secreted proteins will restructure most of their ER into cisternae[108,116].

A notable specialised ER morphology is the arrangement of sheets of smooth ER into tightly stacked arrays. These membrane structures can take various shapes and forms, such as concentrated whorls and regular sinusoidal arrays with cubic symmetry[117,118] (Figure 1.2).

Transition from one morphology into another is possible and was recently shown to involve the endosomal sorting complexes required for transport (ESCRT) in yeast[119]. Collectively these ER morphologies are known as organised smooth endoplasmic reticulum (OSER)[117].

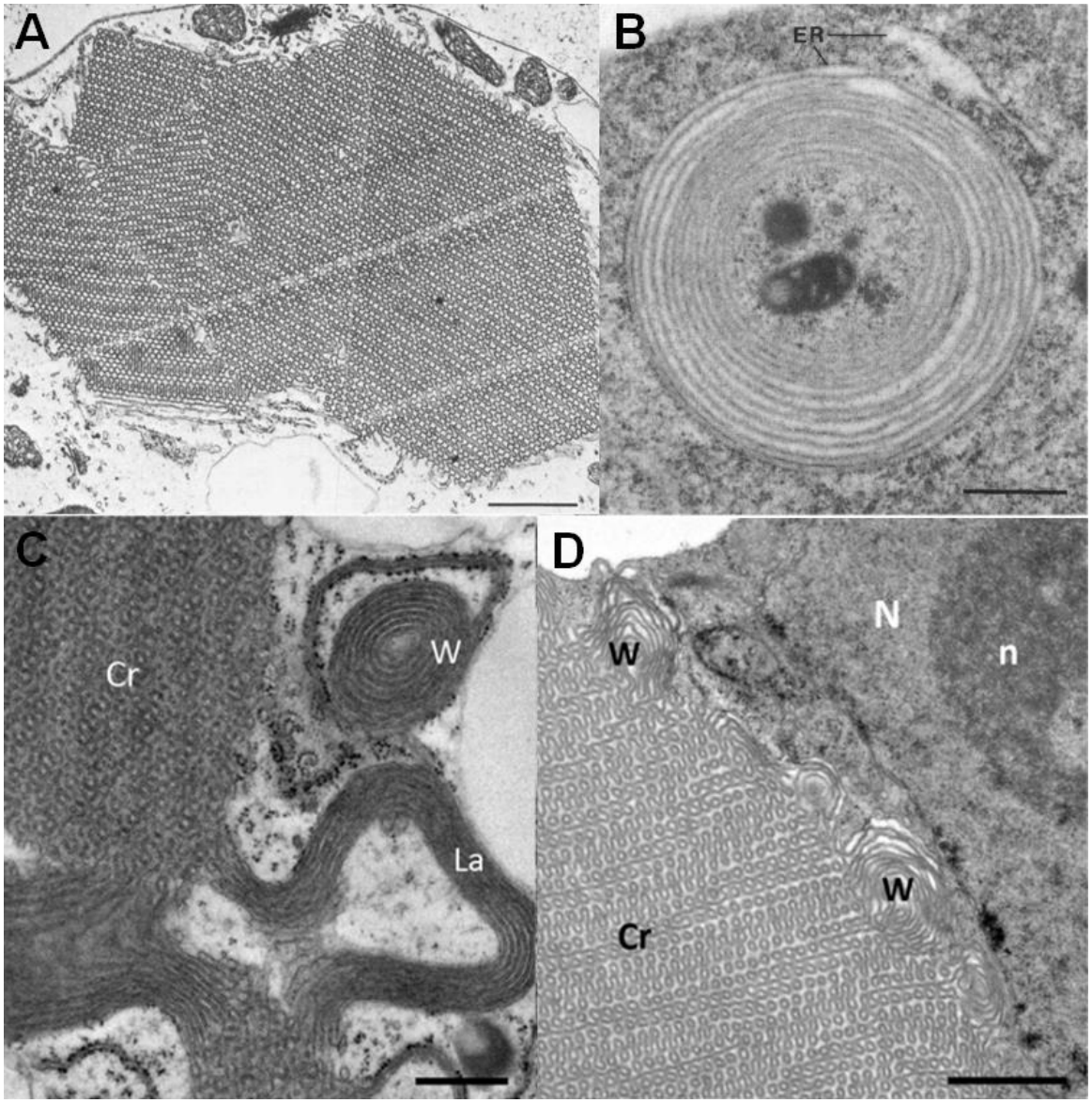


Figure 1.2: Examples of OSER morphologies in plant cells. A: Native crystalloid ER in phloem parenchyma cells of *Helleborus lividus* (figure from Behnke, 1981). B: OSER whorl induced by overexpression of integral membrane M protein of the coronavirus avian infectious bronchitis virus with an oligomerizing ß-glucoronidase attached to the cytoplasmic domain (figure is from Gong *et al.,* 1996). C and D: Crystalloid (Cr), whorl (W) and lamellae (La) OSERs in *N. benthamiana* leaves transiently expressing HMGR membrane domain with a cytoplasmically fused dimerizing GFP. (Figure is from Ferrero *et al.,* 2015). N, nucleus; n, nucleolus. Scale bars: 2.5 µm(A), 0.5 µm (B–D). This figure is re-used from my own previous work, Sandor *et al.* 2021.

OSER structures were first observed with the advent of electron microscopy in highly specialised tissues of animals (e.g., pigment epithelial retina cells[120,121], and neurons[122,123]). OSER is also widespread in plants, especially in the higher plant vasculature where it was described in the sieve elements of the phloem in a range of species, including beans[124], water lilies[125], spinach[126], squash[127], cotton[128], palms[129], elm[130], maples[131] and species of the buttercup family[132]. It was also found in parenchyma cells of willows, where it was hypothesised to be involved in the metabolically dormant state over the winter period[133,134] and similar structures were detected during the pollen germination and tube formation of *Nicotiana alana*[135,136].

Modern tools for recombinant protein expression gave investigations into OSER fresh momentum. A number of synthetic OSER structures have been described in mammalian[117,137–140] and yeast cells[119,141,142]. Plants lost prominence in the study of synthetic OSERs, and most of our understanding of plant synthetic OSERs comes from investigations into the cellular secretory pathway[143]. However, a recurring theme was emerging across all kingdoms: synthetic OSERs were usually observed as an accidental by-product following the overexpression of a recombinant protein eliciting drastic morphological changes in the ER[117,137,142]. The causal recombinant proteins generally consisted of an ER-resident transmembrane domain, with an oligomerising enzyme or fluorescent marker (usually green fluorescent protein (GFP)) genetically fused to the cytosolic terminus of the construct[117,137,138,144,145]. For example, in tobacco the expression of a recombinant fusion protein consisting of the ER integral membrane protein M of the coronavirus avian infectious bronchitis virus and a head-to-head oligomerising ß-glucoronidase enzyme to observe trafficking through the Golgi led to the formation of large

membranous whorls (Figure 1.2B), where the protein construct accumulated. Expression of the ß-glucoronidase alone was not sufficient to induce the formation of OSER[144].

However, there remains significant controversy in the field about the formation of OSER. Plenty of evidence has been gathered about the necessity of a cytosolically-facing head-to-head oligomerising domain to induce OSER, mostly by showing that removal of this domain abolished the OSER-inducing capacity of the recombinant proteins[117,137,139,141,146,147]. However, evidence of the opposite has been discovered as well. For example, in plants, the overexpression of the ER-membrane resident tetramerising 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase (HMGR) caused the formation of OSER, even when the tetramerising domain was completely removed[148,149]. Furthermore, disruption of the ER-to-Golgi protein transport by monomeric proteins without a transmembrane domain was also sufficient to induce OSER in some cases[150,151].

The accumulation of seemingly contradictory experimental evidence led to the development of several competing theories of OSER formation[143]. Some of these emphasised the importance of the oligomerisation domain and the recombinant protein being expressed to sufficiently high levels[117,140]. Others suggested that the transmembrane domain is the key factor for OSER formation by inducing membrane proliferation and erroneous membrane architecture as native ER organising factors (e.g., reticulons) are diluted out[143,148]. A third alternative hypothesis that was raised focused on the disruption of ER-to-Golgi protein transport as another putative causal force, by causing the retention of other OSER-inducing factors[119,143,152]. Recently however it was shown in Sandor *et al.* (2021) that these three competing hypotheses on OSER formation could possibly be integrated into a single unified theory (Figure 1.3).

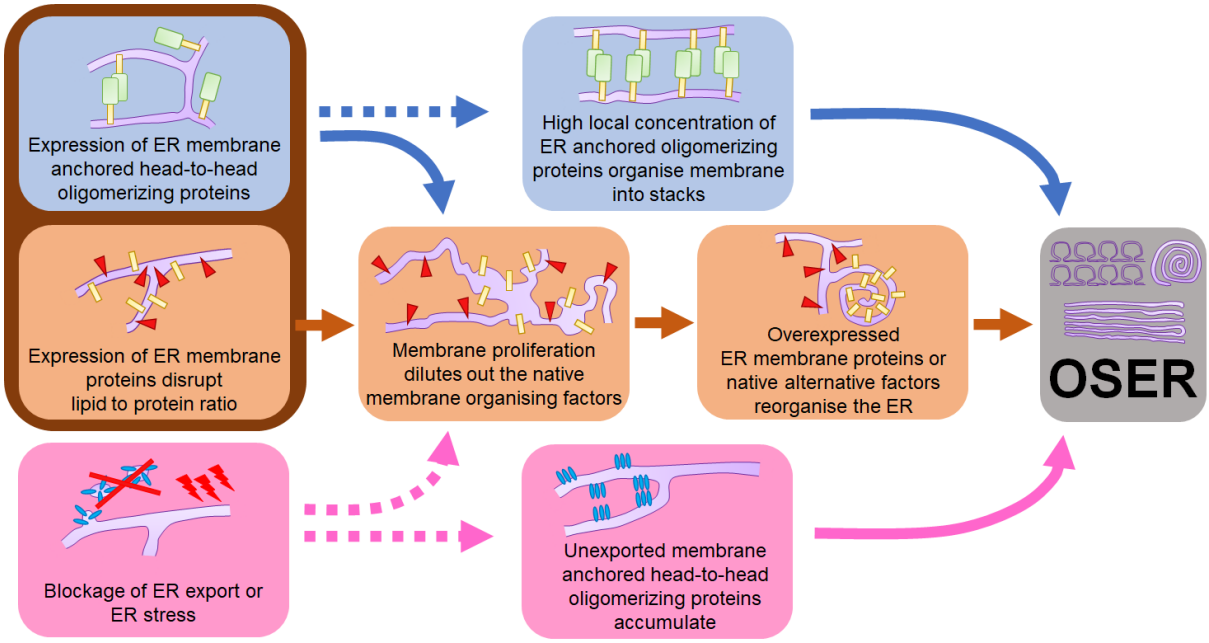


Figure 1.3: A unified theory of OSER formation. Three competing models can be integrated suggesting several alternative pathways that can trigger OSER formation. Models focusing on the sufficiently high expression levels of ER-resident membrane proteins with a head-to-head oligomerising cytosolic domain (blue) could be seen as a special case of its alternative model (orange) which suggest that the primary driving force of OSER formation is membrane proliferation and dilution of native ER organising factors. The possibility of a second parallel pathway (as part of the blue model) cannot be completely excluded. The third pathway (pink) is triggered by the blockage of ER-to-Golgi export, but it is unclear if OSER is formed by the same mechanisms as for the other pathways or is completely independent. This figure is re-used from my own previous work, Sandor *et al.* 2021.

A morphology similar to OSER was discovered by Ian Moore and Marketa Samalova (Department of Plant Sciences, University of Oxford, Oxford, UK, unpublished) after transient expression in mature tobacco leaf epidermis of a synthetic protein formed from an ER-targeted transmembrane domain (the 22 amino acid long BP22 domain derived from the BP80 vacuolar sorting receptor from *Pisum sativum*)[153,154] flanked by two dimerising fluorescent proteins on either side (Figure 1.4). Interestingly, while BP22 was expected to be trafficked through the endomembrane system to the plasma membrane[154], it was integrated into the ER membrane, while the N-terminal GFP was facing the lumen and the

C-terminal yellow fluorescent protein (YFP) the cytosol (due to this organisation, the construct was named G22Y). When expressed transiently, these proteins caused the formation of a single large self-organising *de novo* compartment in transformed leaf epidermal cells. The structure was hypothesised to form as a result of each G22Y polyprotein interacting with a different G22Y polyprotein on either side of the ER membrane, forming an oligomer (Figure 1.4A), blocking the further downstream trafficking of G22Y. The exact process of macro-scale formation of the large compartments was unclear: one possibility raised was a passive mechanism facilitated by the transmembrane domain or the oligomerising domains which caused the aggregation of smaller G22Y "patches" into a larger compartment. Alternatively, an active mechanism mediated by the ER (e.g., a detoxification mechanism) which collects the oligomers was also suggested[155,156].

The structures were further investigated by Nicholas Meuli and Matt White (Department of Plant Sciences, University of Oxford, Oxford, UK) as part of their undergraduate projects under the supervision of Ian Moore[155,156]. When the cytosolic facing GFP was replaced with a monomeric red fluorescent protein (mRFP), the formation of the large compartment was disrupted, and numerous small compartments were observed, suggesting that the dimerization of the cytosolic protein plays a key role in the formation of the single large compartment[155,156]. While this is in line with previous findings about OSER formation[117,143], the structures observed by G22Y overexpression was morphologically different from reported OSER structures and usually only a single compartment was formed, unlike in OSER (Figure 1.4C&D). The experiment replacing GFP with mRFP also showed that the synthetic protein can be modified, allowing replacement of the individual components, such as the dimerising domains. This modularity was further tested by genetically fusing an

mRFP cargo protein to the N-terminus of the G22Y scaffold, which was shown to concentrate into the compartment as expected[155,156].

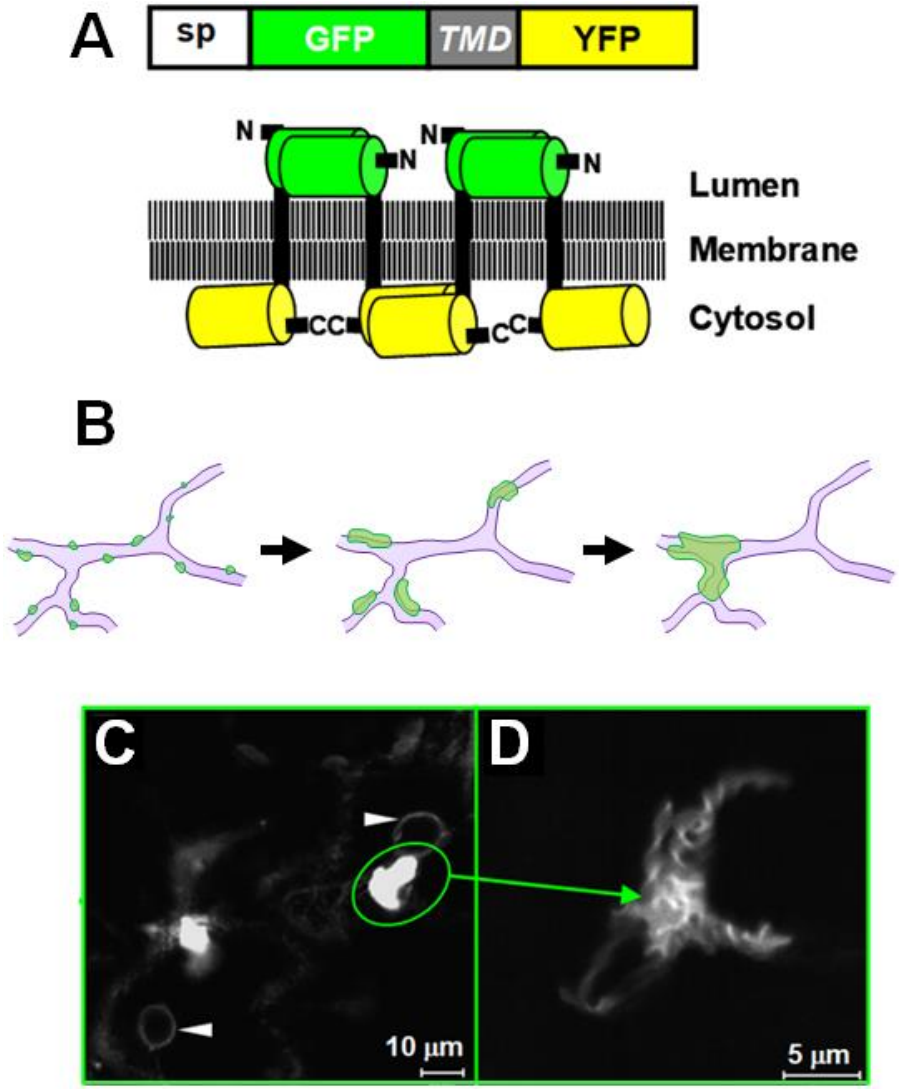


Figure 1.4: Remodelling the ER using the synthetic oligomerising G22Y polyprotein. A: Schematic of the G22Y polyprotein components and its predicted assembly via oligomerisation across the ER membrane as proposed by Moore and Samalova. They hypothesised that the G22Y polyproteins form a long oligomer chain across the ER membrane, causing the formation of a modified ER compartment. sp: ER-targeting signal peptide. TMD: BP22 transmembrane domain. B: Proposed aggregation of smaller G22Y "patches" into the larger compartment structure. C&D: Confocal microscopy images of G22Y compartments, with the GFP signal visualised. G22Y was transiently expressed in adult tobacco leaf epidermis. D was captured using high-resolution airy-scan imaging of the compartment. Arrowheads show the nucleus. Scale bars: 10 µm in C, 5µm in D. Figure A, C and D courtesy of Ian Moore and Marketa Samalova (unpublished).

1.5 Project rationale and aims

The studies carried out by Moore *et al.* were only preliminary in their scope and the discovered compartment warrants further investigation if it is to be exploited for synthetic biology. Since the G22Y compartment formation requires only the expression of a single protein construct, it has great potential to be a simple method of generating a novel synthetic membranous compartment. Such a platform in plants has long been sought after by synthetic biologists and could have considerable impact on the field.

To determine its suitability as a plant synthetic biology platform, in-depth characterisation of the compartment system is required. This would also enable me to determine its relationship to OSER. If it is an independent novel ER morphology, then its characterisation could provide new insight into plant ER structure and plasticity. On the other hand, if this compartment is a subtype of OSER, then a thorough characterisation of it in plants is also valuable, since previously OSER was primarily investigated in yeast and mammalian cell cultures, and significant gaps exist in the plant OSER literature that could be filled[143]. This is especially important when considering that plant ER is not identical in properties to yeast and mammalian ER, and this could have significant effect on OSER as well. For example, in plant cells the large central vacuole causes ER to be sandwiched between the tonoplast and the plasma membrane, unlike in other eukaryotes[107,157].

In general, investigations into OSER have focused on its effects on the host ER and the resultant ER architecture. While the potential of OSER as a novel synthetic membranous compartment has been raised before[31,143], the practical applications of OSER have not been investigated. Furthermore, no comprehensive study has been carried out to date to assess the plasticity or possible modifications to OSER structure. If the G22Y compartments are a

subtype of OSER, this could be attempted due to the amenability of the G22Y protein to replace, add or remove its constitutive domains.

Understanding the impact of the G22Y compartments on the host plant combined with a detailed characterisation of the morphology of the compartments would allow the suitability of this system for potential synthetic biology applications to be assessed. As a synthetic membranous compartment, it could be a valuable tool for molecular farming for improved high-value protein production or to serve as a custom platform for the formation of enzyme-rich microdomains for plant industrial biosynthesis. The full range of possible applications enabled by this compartment system will depend on its properties. Following the characterisation of the G22Y compartment system, proof-of-principle experiments can be carried out to test these putative applications and determine if the compartments discovered by Ian Moore and Marketa Samalova could be competitive with currently available synthetic compartmentalisation strategies and become a popular synthetic biology tool for plants.

In summary, the key goals of this research project are to:

(1) Extend the characterisation of the G22Y compartment system by expanding the available suite of compartment-forming protein constructs.
(2) Determine its relationship to OSER, and if it is a subtype of OSER, expand our understanding of the properties and formation of plant OSER.
(3) Assess the impact of the G22Y compartments on the plant hosts.
(4) Carry out pilot studies on the feasibility of using the compartment for potential applications (such as protein production or as a synthetic microdomain) based on the results of (1) and (3).

Chapter 2: Materials & Methods

2.1 Genetic construct design

All oligomerising constructs were synthesised by Twist Bioscience (San Francisco, USA) and supplied recombined in the Gateway entry vector pTWIST-ENTR, with the exception of the original G22Y construct[155,156], which was a gift from Ian Moore (Department of Plant Sciences, University of Oxford, Oxford, UK). Encoded protein constructs were modelled using pyMOL (v2.3.4, Schrödinger LLC, New York, USA) to determine the binding orientation of the oligomerising domains and to reduce the possibility of structural interference to oligomerisation caused by the genetic fusion of the protein domains used. Plasmids were designed and assembled using SnapGene (v3.0.3, GSL Biotech, San Diego, USA). The individual gene parts used for the design of the constructs can be found in Table 2.1.

Fluorescent markers RFP-HDEL (for the ER lumen)[158], TAR2-RFP (for the ER membrane)[159], YFP-PEX2 (for peroxisomes)[160], GFP-TIP1:1 (for the vacuolar membrane)[158], used for subcellular localisation were courtesy of Verena Kriechbaumer (Department of Biological and Medical Sciences, Oxford Brookes University, Oxford, UK). The cytosolic marker Peredox-mCherry[161] was provided by Beatriz Moreno Garcia (School of Life Sciences, University of Essex, Colchester, UK). The SpyCatcher003 and SpyTag003 constructs[162] were provided by Mark Howarth (Department of Biochemistry, University of Oxford, Oxford, UK).

2.2 Molecular cloning

Polymerase chain reactions (PCR) were carried out using Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, USA) according to the manufacturer's instructions. Primers used for PCRs and sequencing were designed using SnapGene software (GSL Biotech, San Diego, USA) using the default values in the software, and were synthesised by Thermo Fisher Scientific (Waltham, USA). Their sequences and descriptions can be found in Table 2.2. PCR products were subsequently assembled for transformation by using the NEBuilder HiFi DNA Assembly Cloning Kit (New England Biolabs (UK), Hitchin, UK) immediately after PCR, according to the manufacturer's instructions.

Alternatively, some PCR products were assembled via blunt end cloning, as follows: samples were incubated with 10 U DpnI (New England Biolabs (UK), Hitchin, UK) in 1x CutSmart Buffer (New England Biolabs (UK), Hitchin, UK) at 37 °C for 30 minutes to remove any template DNA. This was followed by purification of the PCR products using the QIAquick PCR purification kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The purified samples were phosphorylated by incubating them with 10 U T4 Polynucleotide Kinase (Thermo Fisher Scientific, Waltham, USA) and 1x T4 PNK Buffer (Thermo Fisher Scientific, Waltham, USA) supplemented with 10 mM ATP at 37 °C for 30 minutes, followed by addition of 10 U T4 DNA Ligase (Thermo Fisher Scientific, Waltham, USA) and incubation at room temperature for 30 min to circularise and ligate the plasmids.

Constructs in the Gateway entry plasmid pTWIST-ENTR were subcloned into the plant Gateway expression plasmid pK7WG2[163] using the Gateway LR Clonase II Enzyme Mix (Thermo Fisher Scientific, Waltham, USA) according to the manufacturer's instructions. Following recombination, the plasmid was purified using the QIAquick PCR purification kit

(Qiagen, Hilden, Germany) according to the manufacturer's instructions and then transformed into *E. coli.* For molecular cloning and amplification of plasmids the α-select Silver Chemically Competent Cells (Meridian Bioscience, Cincinnati, USA) and OneShot Mach1 T1 Phage-Resistant Chemically Competent Cells (Thermo Fisher Scientific, Waltham, USA) *E. coli* strains were used. For the transformation 100 ng of plasmid DNA was incubated with the competent cells for 20 min on ice, followed by a 45 sec, 42 °C heat-shock. Cells were then returned to ice for 5 min and then incubated at 37 °C in 500 µL Lysogeny Broth (LB) media (10 g L^{-1} tryptone, 10 g L^{-1} NaCl, 5 g L^{-1} yeast extract) for 60 min before plating on solid LB-agar (LB media supplemented with 15 g L^{-1} agar) in 8.5 cm diameter Petri dishes supplemented with the appropriate antibiotics (50 µg mL^{-1} kanamycin for pTWIST-ENTR and 50 µg mL^{-1} spectinomycin for pK7WG2).

E. coli competent cells were prepared as follows[164]: a 100 mL liquid LB media culture of the desired strain was grown at 37 °C in a shaker until the optical density reached OD_{600} = 0.5, then centrifuged at 4 °C for 10 min at 4500 *g* to pellet the cells. The pellets were resuspended in 33 mL ice-cold RF1 buffer (100 mM RbCl, 50 mM $MnCl_2$, 30 mM Potassium-acetate, 10 mM $CaCl_2$, 15 % (w/v) glycerol, pH = 5.8 adjusted with KOH) and allowed to rest for 60 min at 4 °C. The centrifugation step was repeated (as above), and the pellets were resuspended in 8 mL of ice-cold RF2 buffer (10 mM 3-(N-morpholino)propanesulfonic acid, 10 mM RbCl, 75 mM $CaCl_2$, 15 % (w/v) glycerol, pH = 6.8 adjusted with HCl). After incubation for 15 min at 4 °C, the cells were dispensed into Eppendorf tubes pre-cooled to -80 °C and were immediately frozen in liquid N_2 and stored at -80 °C until used.

To acquire purified plasmid DNA, individual colonies were grown in 10 mL liquid LB media supplemented with the appropriate antibiotics (50 µg mL^{-1} kanamycin for pTWIST-ENTR

and 50 µg mL^{-1} spectinomycin for pK7WG2) overnight in a 37 °C shaker, and the DNA was subsequently extracted using the QIAprep Spin Miniprep Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Prior to any further use, antibiotic resistant colonies were always confirmed to contain the expected constructs using Sanger sequencing (Source Bioscience, Nottingham, UK and Genewiz, Leipzig, Germany). Primers used for the sequencing reaction can be found in Table 2.2.

Confirmed plasmids were then transformed into *A. tumefaciens* (strains LBA4404 and AGL1, courtesy of Beatriz Moreno Garcia (School of Life Sciences, University of Essex, Colchester, UK) and Philippe Varennes-Jutras (Department of Plant Sciences, University of Oxford, Oxford, UK) respectively) as follows: 1000 ng of the plasmid was incubated with the competent cells for 5 min, followed by freezing in liquid nitrogen for 5 min. Samples were thawed at room temperature and then incubated at 28 °C in 500 µL LB media for 4 h before plating on LB-agar plates supplemented with the appropriate antibiotics.

A. tumefaciens competent cells were prepared as follows: a 100 mL liquid LB media culture (supplemented with 50 µg mL^{-1} rifampicin) of the desired strain was grown at 37 °C in a shaker until optical density reached OD_{600} = 0.75, then centrifuged at 4 °C for 10 min at 4500 *g* to pellet the cells. The pellets were resuspended in 10 mL of ice-cold 20 mM $CaCl_2$, then the centrifugation was repeated. The pellets were resuspended in 2 mL of ice-cold storage media (20 mM $CaCl_2$, 10 % (v/v) glycerol), dispensed into Eppendorf tubes pre-cooled to -80 °C and were immediately frozen in liquid N_2 and stored at -80 °C until used.

2.3 Plant growth conditions

N. tabacum and *N. benthamiana* were grown in a greenhouse with a 16 h light – 8 h dark photoperiod at 25 °C, with natural light supplemented to achieve light intensity of up to

200 µM m^{-2} s^{-1} using ATTIS-7 LED grow lights (Plessey, London, UK). Plants were grown in 11 cm square black pots on regularly watered soil composed of a 3:1 ratio of Sinclair Pro modular seed growing compost (Sinclair Pro, Chesire, UK) and Sinclair pro fine vermiculite (Sinclair Pro, Chesire, UK) with 0.4 g L^{-1} Exemptor (ICL, Ipswich, UK) as an insecticide.

Stably transformed *A. thaliana* seeds carrying the C22Y or 22Y constructs (seeds courtesy of Federica Brandizzi, Department of Plant Biology, Michigan State University, East Lansing, USA) were ethanol-sterilised and sowed onto MS-agar (4.33 g L^{-1} Murashige & Skoog medium (Duchefa Biochemie, Haarlem, Netherlands), 15 g L^{-1} agar) plates supplemented with 50 µg mL^{-1} kanamycin for plant selection and 200 µg mL^{-1} cefotaxime to reduce microbial contamination. After stratifying for 3 d in dark at 4 °C, the plates were moved to a 20 °C growth cabinet with a 16 / 8 h light / dark cycle (at light intensity of 120 µM m^{-2} s^{-1}) and surviving seedlings were subsequently transplanted individually onto soil and grown in a greenhouse in identical conditions as described above, with the exception of using 5.5 cm square black pots. Wild-type Col-0 *A. thaliana* were grown in identical conditions, with the exception of excluding kanamycin from the MS-agar plates.

2.4 Plant transient transformation

N. tabacum and *N. benthamiana* plants between 4 and 6 weeks old were transiently transformed by injecting *A. tumefaciens* suspended in infiltration buffer (5 mM 2-(N-morpholino)ethanesulfonic acid, 5 mM $MgCl_2$, 500 µM acetosyringone, 500 µM ascorbic acid, 0.03 % (v/v) Tween20, pH = 5.6 adjusted with HCl)[165] at optical density OD_{600} = 0.5 into the underside of fully expanded mature leaves using a 1 mL plastic needleless syringe. When using AGL1 *A. tumefaciens* strains, these were supplemented with a 1:1 ratio of RNA silencing suppressor p19 carrying strains (courtesy of Philippe Varennes-Jutras

(Department of Plant Sciences, University of Oxford, Oxford, UK)) in the infiltration buffer to repress siRNA mediated RNA silencing of the introduced constructs and thus increase transient expression levels of the target protein[166].

When co-expressing multiple protein constructs, the respective *A. tumefaciens* strains were mixed at equal ratios (normalised by optical density) in the infiltration buffer. In some cases, especially when testing the attachment of cargo proteins to the surface of the compartments, co-expression of multiple protein constructs was done via sequential transformation: in these instances, *A. tumefaciens* strains carrying the plasmids of the compartment forming constructs were introduced as described above, then after 7 d, the infiltration was repeated using the same protocol in the same leaves with *A. tumefaciens* strains carrying the plasmids of the cargo.

Protein expression was routinely tested for each infiltration using confocal microscopy to visualise the fluorescent proteins encoded as part of each construct.

2.5 Phenotypic analysis of stably transformed *A. thaliana* plants

Whole *A. thaliana* seedlings on MS-agar plates were imaged 5 d after moving to a growth cabinet using a Leica M165FC (Leica Microsystems, Wetzlar, Germany) microscope equipped with a 10x air-immersion objective. YFP fluorescence was measured using the YFP filter set, with excitation at 480-520 nm and emission at 505-565 nm, with an approximately 1 sec exposure time, to avoid saturation of the fluorescent signal.

To measure leaf surface area of grown plants, they were photographed from above at 4 and 5-weeks-old using the digital camera of a Xiaomi Redmi 5 (Xiaomi Corporation, Beijing, China) at 3000x4000 px resolution. The collected images were analysed using LeafLab

(v1.5), a custom image analysis software built in Matlab, developed by Mark Fricker (Department of Plant Sciences, University of Oxford, Oxford, UK, unpublished). LeafLab segments the plants from the soil and tray using colour and light intensity thresholds which are manually defined. The basic properties (e.g., surface area, solidity, perimeter length, etc) of the identified plants are then determined automatically by LeafLab. A Nikon D700 digital camera (Nikon Corporation, Tokyo, Japan) at 4256x2832 px resolution was used for side-by-side comparison photographs of grown plants.

To measure dry weight, 7-week-old *A. thaliana* shoots were dried in an oven at 60 °C for 4 d. Samples were weighed regularly over that period and the dry weight recorded once a stable value was achieved.

2.6 Purification of transiently expressed proteins from mature *N. benthamiana* leaves

N. benthamiana was preferentially used for purification experiments due to its improved amenability to higher level expression of transiently introduced constructs than *N. tabacum.* Transformed leaves were harvested 7 d after infiltration and frozen in liquid nitrogen, then thawed at room temperature. Samples were homogenised using a pestle and mortar in 10x volume of grinding buffer (50 mM NaH_2PO_4, 150 mM NaCl, 50 mM ascorbic acid, 0.6 % (w/v) PVPP-40, 0.4 % (w/v) bovine serum albumin, 5 % (v/v) glycerol, 1 % (v/v) Tween20, 1 mini tablet of Pierce Protease Inhibitor (Thermo Fisher Scientific, Waltham, USA), pH = 8.0 adjusted with NaOH) with a pinch of acid-washed sand. The rough lysate was sonicated using a Grant Ultrasonic Bath XB3 (Keison Products, Chelmsford, UK) for 10 min and centrifuged for 15 min at 4 °C at 4000 *g* to remove cell debris, sand and unbroken cells.

For nickel-affinity purification of polyhistidine-tagged proteins, the supernatant was added to a 20 mL Econo-Pac gravity-flow chromatography column (Bio-Rad, Hercules, USA) pre-packed with 4 mL of Ni-NTA Agarose resin (Thermo Fisher Scientific, Waltham, USA). The samples were allowed to bind the resin for 18 h on a rotating shaker at 4 °C, then washed 5x with 10 mL wash buffer (50 mM NaH_2PO_4, 150 mM NaCl, 10 mM Imidazole, 5 % (v/v) glycerol, 1 % (v/v) Tween20, pH = 8.0 adjusted with NaOH). The purified protein samples were eluted using 3 washes of 3 mL of elution buffer (50 mM NaH_2PO_4, 500 mM NaCl, 250 mM Imidazole, 5 % (v/v) glycerol, 0.1 % (v/v) Tween20, pH = 8.0 adjusted with NaOH).

For purification of proteins by immunoprecipitation of mCherry, the supernatant from the crude lysate was added to a ChromoTek (München, Germany) 500 µL spin column packed with 25 µL of RFP-trap agarose bead slurry (ChromoTek, München, Germany). The samples were allowed to bind to the RFP-trap for 60 min on an end-over-end tumbler at 4 °C, then washed 3x with 500 µL of wash buffer (10 mM Tris(hydroxymethyl)aminomethane (Tris)-HCl (pH = 7.5), 150 mM NaCl, 0.5 mM EDTA, 5 % (v/v) glycerol, 1 % (v/v) Tween20). The purified samples were eluted using 2x 50 µL 200 mM glycine (pH = 2.5 adjusted with HCl) into a 2 mL Eppendorf tube containing 10 µL of 1 M Tris base (pH = 10.4).

2.7 Fluorescent protein quantification

100 µL of each sample from each step of the protein purification was analysed in a DTX880 Multimodal microplate reader (Beckman Coulter, Brea, USA). Samples were placed in a Nunc 96-well polystyrene round bottom microwell plate (Thermo Fisher Scientific, Waltham, USA). Fluorescent measurements for YFP were taken using a 465-505 nm excitation filter and a 510-560 nm emission filter. Each sample was measured three times

to account for technical variation. The output was averaged for each sample and normalised for background levels.

2.8 Analysis of purified proteins by sodium-dodecyl-sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

Gels were cast using the SureCast Gel Handcast System (Thermo Fisher Scientific, Waltham, USA), according to the manufacturer's instructions. 8 mL of resolving gel was used for each plate, composed of 2 mL of 4x SureCast Resolving Buffer (1.5 M Tris-HCl (pH = 8.8), 0.4 % (w/v) SDS), 2 mL of 40 % (w/v) acrylamide solution (with a 29:1 acrylamide:bis-acrylamide ratio), 3.9 mL of deionized water, 90 µL of freshly prepared 10 % (w/v) ammonium persulphate solution (APS) and 10 µL of N, N, N', N' Tetramethylethylene-1,2-diamine (TEMED) to trigger polymerisation of the gel. The resolving gel mixture was immediately loaded into the SureCast Gel Handcast System and covered with 1 mL of isopropanol. Once the gel was set, the isopropanol was removed and 4 mL of stacking gel was prepared using 0.75 mL of 4x SureCast Stacking Buffer (0.5 M Tris-HCl (pH = 6.8), 0.4 % (w/v) SDS), 0.3 mL of 40 % (w/v) acrylamide solution (with a 29:1 acrylamide:bis-acrylamide ratio), 1.9 mL of deionized water, 40 µL of freshly prepared 10 % (w/v) APS and 5 µL of TEMED to trigger polymerisation of the gel. The stacking gel was covered with a 10-, 12-, or 15-well comb.

Samples were mixed with SDS-PAGE loading buffer (50 mM Tris-HCl (pH = 6.8), 2 % (w/v) SDS, 6 % (v/v) glycerol, 2 mM DTT, 0.01 % (w/v) Bromophenol Blue) and incubated for 30 min at 37 °C before loading into the wells. The gels were run for 2 h at 120 V constant voltage in SDS-PAGE running buffer (25 mM Tris base, 200 mM glycine, 3.5 mM SDS) until the Bromophenol dye front had reached the bottom of the gel. Gels were then used for Western Blotting. Some SDS-PAGE gels were also fluorescently imaged directly using an

ImageQuant LAS4000 camera system (GE Healthcare, Illinois, USA) using a 520 nm LED light for excitation and a 555-595 nm filter for emission detection for YFP.

2.9 Western Blotting

After SDS-PAGE, protein bands were transferred from the acrylamide gel to a nitrocellulose membrane using electroblotting (100 V, 60 min) in ice-cold transfer buffer (20 % (v/v) methanol, 50 mM Tris base, 250 mM glycine, 3.5 mM SDS). The nitrocellulose sheet was blocked using milk albumin (50 mM Tris-HCl (pH = 7.5), 150 mM NaCl, 0.5 % (v/v) Tween20, 8 % (w/v) skimmed milk powder) for 60 min at room temperature on a shaker, then incubated with the primary antibody (1 µg mL^{-1}) overnight at room temperature on a shaker. Polyclonal rabbit IgG anti-mCherry and anti-GFP primary antibodies (Proteintech Europe, Manchester, UK) were used for protein constructs containing mCherry and YFP, respectively. The membrane was subsequently washed 3x with Tris-buffered saline with Tween20 (TBS-T: 50 mM Tris-HCl (pH = 7.5), 150 mM NaCl, 0.5 % (v/v) Tween20) for 5 min each, then incubated with the secondary antibody (0.1 µg mL^{-1} polyclonal goat anti-rabbit IgG with a conjugated horseradish peroxidase (HRP - Sigma-Aldrich, St. Louis, USA)) for 60 min at room temperature on a shaker, then washed 3x with TBS-T for 5 min each. Before visualisation, the membrane was treated for 2 min with 0.1 mL cm^{-2} of EZ-ECL Chemiluminescence Detection Kit for HRP (Biological Industries, Kibbutz Beit-Haemek, Israel), composed of a 1:1 ratio of peroxide solution and luminol. The HRP signal was visualised using a chemiluminescence detector ImageQuant LAS4000 (GE Healthcare, Illinois, USA).

2.10 Confocal laser scanning microscopy

Leaf epidermal samples of *A. thaliana, N. tabacum* and *N. benthamiana,* and seeds and stem samples of *A. thaliana* were imaged using a Zeiss LSM880 confocal microscope equipped with an airy-scan detector (Carl Zeiss AG, Oberkochen, Germany). Images were commonly captured with a C-Apochromat 40x/1.2 W autocorrect M27 water-immersion objective (Carl Zeiss AG, Oberkochen, Germany) at 1024x1024 px resolution at pixel spacing of 20-100 nm with 4-line averaging and 16-bit depth at maximum speed, with typical excitation at 458 nm (CFP), 488 nm (GFP), 514 nm (YFP, fluorescein diacetate (FDA) and chloroplast) or 561 nm (RFP and mCherry) and emission at 460-480 nm (CFP) 490-510 nm (GFP), 520-560 nm (YFP, FDA), 590-640 nm (RFP and mCherry) or 660-700 nm (chloroplasts). When imaging two fluorophores in one sample, care was taken to avoid spectral overlap by using sequential excitation and line switching.

For experiments focusing on quantitative analysis of the ER network and dynamics and fluorescence recovery after photobleaching (FRAP), the delay between the images in time-series was minimised by reducing image resolution to 512x512 px and using 2-line averaging to give a frame rate of approximately 125 ms image^{-1}. For FRAP experiments, leaf epidermal samples were treated with 25 µM Latrunculin B (Sigma-Aldrich, St. Louis, USA) and 2 µg mL^{-1} FDA (only when imaging cytosolic recovery rate – Sigma-Aldrich, St. Louis, USA) for 10 min at room temperature to immobilise the ER immediately prior to imaging. When determining what proportion of the scaffold protein constructs or cargo proteins were localised to the synthetic ER compartment, whole cell Z-stack images were captured with extra care to avoid any signal saturation.

Images were analysed in ImageJ (v1.52k, U.S. National Institutes of Health, Bethesda, USA) and statistical analysis (where appropriate) was carried out using R (v3.6.1, R Foundation for Statistical Computing, Vienna, Austria) and Microsoft Excel. To determine what proportion of scaffold or cargo proteins were in the compartments, for the appropriate fluorophore the total signal value in the compartments and the total cellular signal value were measured. Then, the total compartment signal values were divided by the total cellular signal value after subtracting background fluorescence. All confocal microscopy images shown in this thesis are optical sections.

2.11 Quantitative analysis of ER network architecture and dynamics

Short time-series airy-scan images (20 sec) were collected of the ER network at the apical end of cells expressing a compartment-forming protein construct and RFP-HDEL. Cells were only imaged if the compartment was not visible, to visualise only the unmodified ER in the region-of-interest. Images underwent quality control in ImageJ (v1.52k) to remove time-series that had drifted in the Z-axis or were blurry. Images drifting in the X or Y axes were corrected (if possible) using the StackReg (v1.0) module[167].

The images were then processed and analysed using the AnalyzER software package (v1.1)[168], following the protocols described in Pain *et al.* (2019). The parameters investigated are described in Table 3.1. These parameters were then statistically analysed using principal component analysis (PCA) using R (v3.6.1, R Foundation for Statistical Computing, Vienna, Austria) to determine if there were statistically discriminating collection of features (kinetic or morphological, in the tubules, cisternae or enclosed areas) between the ER of different compartment-expressing cells and cells only expressing RFP-

HDEL. Using a Scree plot, the top principal components (PC) which cumulatively explain 90 % of variance in the data were selected for multivariate analysis of variance (MANOVA) and analysis of variance (ANOVA). Any PC where the ANOVA detected a significant difference between either of the groups were further investigated using Tukey's HSD test to determine if the difference is between the wild type and any of the compartment expressing groups. If a significant difference was found, the parameters contributing the most to the relevant PC were individually tested to determine if they alone would be significantly different between groups.

2.12 Transmission electron microscopy (TEM)

Mature *N. tabacum* leaf sections were fixed, stained using zinc-iodine-osmium[169] and embedded in resin as described in Kittelmann (2018) by Clare Jones (Department of Biological and Medical Sciences, Oxford Brookes University, Oxford, UK). 90 nm sections were cut from the embedded samples, placed onto 200 mesh copper grids and then post-stained for 5 min with lead-citrate. The stained sections were imaged using a FEI Tecnai T12 transmission electron microscope (FEI, Hillsboro, USA), operated at 120 kW. Images were captured by a GATAN OneView digital camera (AMETEK, Pleasanton, USA) and analysed in ImageJ (v1.52k). TEM sectioning and imaging was carried out by Errin Johnson (Sir William Dunn School of Pathology, University of Oxford, Oxford, UK) as a service.

2.13 Gas chromatography-mass-spectrometry (GC-MS) metabolite identification and quantification

Leaf tissue was extracted and prepared for GC-MS following the protocol outlined in Lisec *et al.* (2006), with some modifications as described below[170]. Leaf material from mature *N. benthamiana* was harvested, weighed (a maximum 200 mg of fresh weight was used per

sample) and immediately frozen in liquid nitrogen. The samples were homogenised in a pre-cooled ball mill for 3 min at 20 Hz. 1.4 mL of 100 % methanol pre-cooled to -20 °C and 50 µL of 100 µg mL^{-1} ribitol (as an internal standard) was added, and the samples were incubated at 70 °C while shaking at 700 rpm for 10 min. The samples were subsequently centrifuged for 10 min at 12000 *g* to pellet any remaining cell debris, the supernatant was moved to a glass vial and 750 µL of pre-cooled chloroform and 1.5 mL water was added. The samples were vortexed and centrifuged for 15 min at 2200 *g*. 1 mL of the upper phase was moved to a fresh Eppendorf tube and dried in a vacuum concentrator at room temperature.

To derivatize the samples, each one was incubated for 2 h with 40 µL of 20 mg mL^{-1} methoxyamine hydrochloride in pyridine at 37 °C while shaking. Subsequently, 60 µL of 1 g mL^{-1} N-Methyl-N-(trimethylsilyl)trifluoroacetamide was added and incubated for 30 min at 37 °C while shaking.

0.5 µL of each derivatised sample was injected in a 1:10 split mode at 230 °C to an Intuvo 9000 gas chromatography-mass spectrometer (Agilent Technologies, Santa Clara, USA) fitted with a 15 m x 250 µm x 0.25 µm DB-5ms Ultra Inert Intuvo GC capillary column module (Agilent Technologies, Santa Clara, USA). Chromatography was achieved using the following temperature programme: 70 °C for 5 min and then heated at a rate of 7.5 °C per min to 330 °C for a total run time of approximately 40 min. An Agilent 5977B mass spectrometer (Agilent Technologies, Santa Clara, USA) was used in scanning mode with electron ionization to record mass ranges between *m/z* 50 to *m/z* 600 after a solvent delay of 7.5 min. Extracted peaks and fragmentation patterns were identified using AMDIS software (v2.7.3) and the National Institute of Standards and Technology (NIST,

Gaithersburg, USA) mass spectrum library (v2.0). After positive identification of their peaks, citramalate and ribitol were quantified using El-MAVEN (v0.11, Elucidata, Cambridge, USA). A citramalate:ribitol standard curve was generated (Supplementary Figure S1) using a pure citramalate standard (Sigma-Aldrich, St. Louis, USA) to allow absolute quantification of citramalate in the samples, followed by statistical analysis in R (v3.6.1, R Foundation for Statistical Computing, Vienna, Austria).

Name of gene or fragment	Description	Sequence or accession number	Reference number (for sequences) or Database (for accession numbers)
GFP	Anti-parallelly dimerising green fluorescent protein mGFP5	U87973.1	European Nucleotide Archive
YFP	Anti-parallelly dimerising yellow fluorescent protein eYFP.	AAX97736.1	GenBank
BP22	Transmembrane domain derived from the BP80 vacuolar sorting receptor of *P. sativum*	TTTTGGCTTGCTCTCGTCGTTTTGATcGCCTTAGCTATGATTGCAGGAGGGGGATTCCTTGTGTAT	[154]
ER localisation signal	ER-targeting signal peptide from the original G22Y construct.	AAGACTAATCTTTTTCTCTTTCTCATCTTTTCACTTCTCCTATCATTATCCTCGGCCGAGTTC	[155]
CC-Di	Parallelly dimerising synthetic coiled coil domain	gagattgccgcgctgaaacaggagattgcggcgctgaagaaagagaacgcagcgctgaaatgggagattgctgcgttgaaacag	[171]
CC-Cap	Antiparallelly dimerising coiled coil domain from *Bos taurus* IF(1) protein.	gcgctgaaaaaacatcatgaaaacgaaattagccatcatgcgaaagaaattgaacgcctgcagaaagaaattgaacgccataaacagagcattaaaaaactgaaacagagcgaagatgatgat	[172]
FDH	Anti-parallelly dimerising enzyme formate dehydrogenase from *Solanum tuberosum* that catalyses the breakdown of formate into CO2. The presequence was truncated to abolish mitochondrial localisation.	CAA79702.2	European Nucleotide Archive
CC-Tet	Parallelly tetramerising synthetic coiled coil domain	GCTGGAGAACTTGCTGCTATTAAGCAAGAACTTGCTGCTATTAAGAAGGAACTTGCTGCTATTAAGTGGGAACTTGCTGCTATTAAGCAAGGAGCTGGAGCT	[171]
CC-NHR2	Anti-parallelly tetramerising coiled coil domain from the *Mus musculus* AML1/ETO chimeric protein.	ctaacagacagagaatgggcagaagagtggaaacatcttgaccatctgttaaactgcataatggacatggtagaaaaaacaaggcgatctctcaccgtactaaggcggtgtcaagaagcagaccgggaagaattgaattactggatccggcggtacagtgacgccgag	[173]
mCherry	Monomeric red fluorescent protein	AY678264.1	GenBank
SpyT	SpyTag003 Peptide capable of forming a covalent bond *in vivo* with SpyCatcher003.	cgtggcgtgcctcatatcgtgatggtggacgcctacaagcgttacaag	[162]

SpyC	SpyCatcher003 Peptide capable of forming a covalent bond *in vivo* with SpyTag003.	GTAACCACCTTATCAGGTTTATC AGGTGAGCAAGGTCCGTCCGGT GATATGACAACTGAAGAAGATA GTGCTACCCATATTAAATTCTCA AAACGTGATGAGGACGGCCGT GAGTTAGCTGGTGCAACTATGG AGTTGCGTGATTCATCTGGTAA AACTATTAGTACATGGATTTCAG ATGGACATGTGAAGGATTTCTA CCTGTATCCAGGAAAATATACA TTTGTCGAAACCGCAGCACCAG ACGGTTATGAGGTAGCAACTCC AATTGAATTTACAGTTAATGAG GACGGTCAGGTTACTGTAGATG GTGAAGCAACTGAAGGTGACG CTCATACTGGATCCAGTGGTAG C	[162]
His	6x Polyhistidine tag enabling purification of the constructs.	CACCATCACCATCACCAT	
N-Linker	Linker sequence between BP22 and the ER facing oligomerising domain. Derived from the original G22Y construct.	TCGACTTGGGCTGCT	[155]
C-Linker	Linker sequence between BP22 and the cytosolic facing oligomerising domain. Derived from the original G22Y construct.	AAATATAGAATTAGGggaggagct ggaggtgctggctcgaccatg	[155]
SL	Long stiff linker between YFP and SpyTag in C22Y-SL-SpyT	GCACCGGCGCCTGCGCCTGCAC CCGCTCCA	[174]
LL	Long flexible linker replacing the C-Linker in 22LLY-SpyT-his.	agcaacagcgaagtgaccaccattctga ccgaaagcccggaatttagcggcggcgg cggcagcggcggcggcggcagcggcgg cggcggcgaactggatgaagaaagcatt ctgaaaaacgaa	[174,175]
Luc	Luciferase enzyme from the firefly *Photinus pyralis*.	AAA29795.1	European Nucleotide Archive
HbpA	2-hydroxybiphenyl-3-monooxygenase from the 2-2'-biphenol breakdown pathway of *Pseudomonas nitroreducens.*	U73900.1	GenBank
HbpC	Extradiol ring cleavage dioxygenase from the 2-2'-biphenol breakdown pathway of *Pseudomonas nitroreducens.*	U73900.1	GenBank
MdCMS	Citramalate synthase enzyme MdCMS_1 from *Malus* x *domestica.*	MW528402	[176]

Table 2.1: Gene parts used for the design of the genetic constructs.

Primer name	Nucleotide sequence	Description
Gibson cloning		
AS_gib_109F	CATTATCCTCGGCCGAATTCTCGACTTGGGCTGCTTTTTG	Primers to generate pTWIST-ENTR 22Y from pTWIST-ENTR C22Y
AS_gib_109R	CAAAAAGCAGCCCAAGTCGAGAATTCGGCCGAGGATAATG	
AS_gib_110F	GAGGTGCTGGCTCGACCATGGAGCAGAAACTTATCTCTGAGGAGG	Primers to generate pTWIST-ENTR G22 from pTWIST-ENTR G22C
AS_gib_110R	TCAGAGATAAGTTTCTGCTCCATGGTCGAGCCAGCACCTC	
AS_gib_122F	ATGGCTAGCATAGACGAAGACCGCTCTTTATTAATGAGTCGTGTAGCTTCTACAGCAG	Replace mC-C22 with a tonoplast targeting sequence for FDH-his
AS_gib_122R	TAATAAAGAGCGGTCTTCGTCTATGCTAGCCATAGCAGTGAGACCAATAGCCTG	
AS_gib_128F	acgaactgtacaagggcgcaATGAGTCGTGTAGCTTCTAC	Binds pTWIST FDH-his(ton) to build pTWIST mCherry-FDH-his
AS_gib_128R	tcctcgcccttgctGaccatAGTGAGACCAATAGCCTGCT	
AS_gib_129F	AGCAGGCTATTGGTCTCACTatggtCagcaagggcgagga	Binds pTWIST Lucdeli-mC-SpC to build pTWIST mCHerry-FDH-his
AS_gib_129R	GTAGAAGCTACACGACTCATtgcgcccttgtacagttcgt	
AS_gib_135F	CGGCGTAAATTACTCGTTCggagctacggccagtgc	Binds Lucdeli-mC-SpC to build pTWIST OmpG-mC-SpC-his
AS_gib_135R	CAGGGTAATAACTTTTTCATAGCAGTGAGACCAATAGCCTG	
AS_gib_136F	AGGCTATTGGTCTCACTGCTATGAAAAAGTTATTACCCTGTACCGCAC	Binds pTWIST C22Y-OmpG-his to build pTWIST OmpG-mC-SpC-his
AS_gib_136R	ccagcactggccgtagctccGAACGAGTAATTTACGCCGACAC	
AS_gib_137F	TGGTTTATCAGGTGGCCATTggagctacggccagtgc	Binds Lucdeli-mC-SpC to build pTWIST MdCMS-mC-SpC-his
AS_gib_137R	TTTTCTGCTGAGAAGGCCATAGCAGTGAGACCAATAGCC	
AS_gib_138F	AGGCTATTGGTCTCACTGCTATGGCCTTCTCAGCAGAAAATCTTC	Binds pTWIST C22Y-MdCMS to build pTWIST MdCMS-mC-SpC-his
AS_gib_138R	ccagcactggccgtagctccAATGGCCACCTGATAAACCAATGA	
mCherry_C22Y_fwd	ctcgaccatgATGGTGAGTAAGGGTGAGGAGGATAAC	Primers used to replace YFP with mCherry in pTWIST-ENTR CCTet-22Y and pTWIST-ENTR NHR2-22Y. (Binds to pTWIST-ENTR Luc-mC-SpC) Courtesy of Anya Lindstrom-Battle
mCherry_C22Y_rev	gtttctgctcCTTGTACAGTTCGTCCATGCCGC	
Backbone_C22Y_fwd	actgtacaagGAGCAGAAACTTATCTCTGAGGAGG	Primers used to replace YFP with mCherry in pTWIST-ENTR CCTet-22Y and pTWIST-ENTR NHR2-22Y. (Binds to pTWIST-ENTR CCTet-22Y and pTWIST-ENTR NHR2-22Y)
Backbone_C22Y_rev	tactcaccatCATGGTCGAGCCAGCACC	

		Courtesy of Anya Lindstrom Battle
mCherry_G22C_fwd	ggccgaattcatggtgagcaagggcgag	Primers used to replace GFP with mCherry in pTWIST-ENTR G22-CCTet and pTWIST-ENTR G22-NHR2. (Binds to pTWIST-ENTR Luc-mC-SpC) Courtesy of Anya Lindstrom-Battle
mCherry_G22C_rev	cccaagttgaTTTGTACAGTTCGTCCATGC	
Backbone_G22C_rev	tgctcaccatgaattcggccgaggataatg	Primers used to replace GFP with mCherry in pTWIST-ENTR G22-CCTet and pTWIST-ENTR G22-NHR2. (Binds to pTWIST-ENTR G22-CCTet and pTWIST-ENTR G22-NHR2) Courtesy of Anya Lindstrom-Battle
Backbone_G22C_fwd	actgtacaaaTCAACTTGGGCTGCTTTTTG	
Blunt end cloning		
AS_BLC_203F	AGTAAAGGAGAAGAACTTTTCAC	Primers to remove ER localisation signal from pTWIST-ENTR G22Y to generate pTWIST-ENTR G-Y(cyt)
AS_BLC_203R	CATTGTGGACTAGTCTCGAC	
AS_BLC_204F	ggctcgaccatgagcaag	Primers to remove BP22 transmembrane domain from pTWIST-ENTR G22Y to generate pTWIST-ENTR G-Y(cyt)
AS_BLC_204R	AGCCCAAGTCGACTTGTATA	
AS_BLC_201F	TCGACTTGGGCTGCTTTTTG	Primers to remove CC-Di from pTWIST-ENTR mCherry-C22-FDH.
AS_BLC_205R	cgtagctcccttgtacagtt	
AS_BLC_208F	gcaggaGAGCAGAAACTTATCTCTGAGGAG	Primers to remove SpyTag from pTWIST-ENTR C22Y-SpyT-his.
AS_BLC_208R	tcctgcggcggcggtca	
AS_BLC_208R_b	tcctgcGCGATACTGAGGTGCC	Primers to remove SpyTag from pTWIST-ENTR mC-C22FDH (with AS_BLC_208F).
AS_BLC_209F	tacgcggaatacttcgaaatg	Remove intron in Luciferase (Luc-mC-SpC)
AS_BLC_209R	cgtGATATTTACTTCAATGTGAGC	
AS_BLC_211F	TAGAGACCTATACCCAGCTT	Primers to remove the his-tag from pTWIST-ENTR C22Y-SpyT-his
AS_BLC_211R	TCTAGACAAATCCTCCTCAGA	
AS_BLC_213F	gcaggacgtggcgtgcc	Primers to remove FDH from pTWIST-ENTR mC-C22FDH-SpyT-his
AS_BLC_213R	catggtcgagccagcacctc	

AS_BLC_215F	gtggacgcctacaagcgttacaagTCTAGATAGAGACCTATACCCAGCTTTC	Primers to introduce SpyTag at the C-terminus of C22Y without a linker to generate C22Y-NL-SpyT.
AS_BLC_215R	catcacgatatgaggcacgccacgCAAATCCTCCTCAGAGATAAGTTTC	
AS_BLC_217F	GCGGCCAAGGAAGCAGCAGCGAAAcgtggcgtgcctcatatcgt	Primers to introduce a stiff linker between the Myc and SpyTag in C22Y-NL-SpyT to build C22Y-SL-SpyT.
AS_BLC_217R	AGCCTCCTTTGCAGCTGCCTCCAAATCCTCCTCAGAGATAAGTTTCTGC	
AS_BLC_219F	tagAGACCTATACCCAGCTT	Remove his-tag from Luc-deli-mCherry-SpyC
AS_BLC_219R	TCTAGAgctaccactggatc	
AS_BLC_220F	GACCGCTCTTTATTAATGAGTCGTGTAGCTTCTACAGCAG	Replace mC-C22 with a tonoplast targetting sequence for FDH-his
AS_BLC_220R	TTCGTCTATGCTAGCCATAGCAGTGAGACCAATAGCCTG	
AS_BLC_225F	atggtgagcaagggcgagg	Primers to remove tonoplast targeting sequence from mC-FDH-his(ton)
AS_BLC_225R	AGCAGTGAGACCAATAGCCTGCT	
Sequencing primers		
AS_seq_05R	TTTGTATAGTTCATCCATGCC	Reverse sequencing primer for GFP in G22Y
AS_seq_06F	ggagtacaactacaacagcca	Forward sequencing primer for YFP and BP22 in G22Y
AS_seq_06R	CCCTCAGGCATGGCGCTCTTGAA	Reverse sequencing primer for mGFP5 + BP22 in G22Y
AS_seq_07F	GCAAAAGAACGGCATCAAAG	Sequencing primers for BP22 in G22Y
AS_seq_07R	gtccagctcgaccaggatgg	
AS_seq_08F	TACAAGTCGACTTGGGCTGC	
AS_seq_08R	cggacacgctgaacttgtgg	
AS_seq_09F	GGAAGTTCATTTCATTCGGAGA	Sequencing primer for G22Y binding upstream of the CaMV35S promoter
AS_seq_09R	CATAAATAACGTCATGCATTACATGT	Sequencing primer for G22Y binding in the terminator
AS_seq_010F	ATGAAGACTAATCTTTTTCTC	Sequencing primer binding at the ER localisation signal
AS_seq_010R	CTATCTAGACAAATCCTCCT	Sequencing primer binding at the Myc
AS_seq_09F_alt	GGAAGTTCATTTCATTTGGAGA	Sequencing primer for pK7WG2 binding in the p35S

AS_seq_011R	aggtcactggattttggttttag	Sequencing primer for pK7WG2 binding in the T35S
AS_seq_M13 F	GTAAAACGACGGCCAGT	Common sequencing primer M13 forward for pTWIST-ENTR
AS_seq_M13 R	CAGGAAACAGCTATGAC	Common sequencing primer M13 reverse for pTWIST-ENTR

Table 2.2: PCR and sequencing primers used to generate constructs.

Chapter 3: Characterisation of a synthetic, self-assembling de novo compartment in plant cells

3.1 Introduction

Initial studies of the compartment discovered by Ian Moore and Marketa Samalova (Department of Plant Sciences, University of Oxford, UK) suggested it to be similar to the OSER structures, but markedly different in morphology. However, since OSER can also take different forms[143], the precise relationship between OSER and the G22Y compartment remained uncertain, and its clarification required further characterisation of the G22Y compartment.

The preliminary observations also left many aspects of G22Y compartment formation unresolved. For example, it was unclear if the compartment was still connected to the ER, although N-glycans on the G22Y proteins remained endo-H sensitive, suggesting that G22Y does not enter the Golgi[155]. The assembly of the compartment also remained unclear, as did the internal structure.

Importantly, short-term transient expression of the G22Y protein did not elicit any abnormal phenotypes in the tobacco epidermal cells (other than the compartment formation)[155,156], probably due to the fact that eukaryotic cells adapt well to even substantial remodelling of the ER[109]. However, as transient experiments only followed the compartment for a few days, further experiments were necessary to verify this claim.

In this chapter I describe work to verify and expand on the results of Moore and Samalova. The G22Y construct was used as the starting point to investigate the possibility of remodelling the ER to generate synthetic, self-organising, *de novo* compartments. As the

modularity of the G22Y polyprotein was already established, a range of new constructs were generated to elucidate the principles of the relationship between protein structure and compartment properties, to enable rational design of this new compartment. To facilitate the rapid screening of these new constructs, transient expression was used. The original G22Y compartment and its derivatives were further characterised to fill in the gaps in our current knowledge, including the internal structure of the compartments and their relationship to the ER to determine if they are still connected. Finally, I comprehensively assessed the impact the new compartments have on the structure and dynamics of the rest of the ER network.

3.2 Results

3.2.1 Design of modular compartment forming protein variants

To establish the primary driving forces behind the formation of the G22Y compartment and to expand the toolkit of compartment forming constructs, five other synthetic polyproteins were designed and named following the original nomenclature: G22, 22Y, G22C, C22Y and G-Y(cyt) (Figure 3.1). G22 and 22Y were generated to assess the requirement for having dimerising proteins on both sides of the membrane by removing either the N-terminal GFP (to form 22Y) or the C-terminal YFP (to form G22). In G22C and C22Y, one of the dimerising fluorescent proteins (YFP in the former, and GFP in the latter) were replaced by the synthetic dimerising coiled-coil domain CC-Di[171], which is only 3 kDa, but of similar length as the more compactly structured (ß-barrelled) GFP and YFP[177]. CC-Di also has a much stronger binding affinity ($K_D < 10^{-8}$ M) and dimerises in parallel[171], in contrast with GFP and YFP which weakly dimerises in an antiparallel fashion ($K_D \approx 10^{-4}$ M)[178]. This allowed us to explore the effects of a different binding structure (in terms of both strength and orientation of the binding proteins) and to create a more compact protein construct. Another benefit is the removal of one of the fluorescent proteins, which reduced spectral overlap for later imaging experiments where co-expression with other fluorescent probes were necessary.

G-Y(cyt) was derived from the original G22Y by removing the transmembrane domain BP22 (leaving only the linker sequences between GFP and YFP) and the ER-localisation signal, to generate a construct that would remain in the cytosol without localising to a membrane. This was done to determine if the presence of membrane is required to form

compartments or protein-based phase-separation forces would be sufficient to elicit the observed G22Y structures.

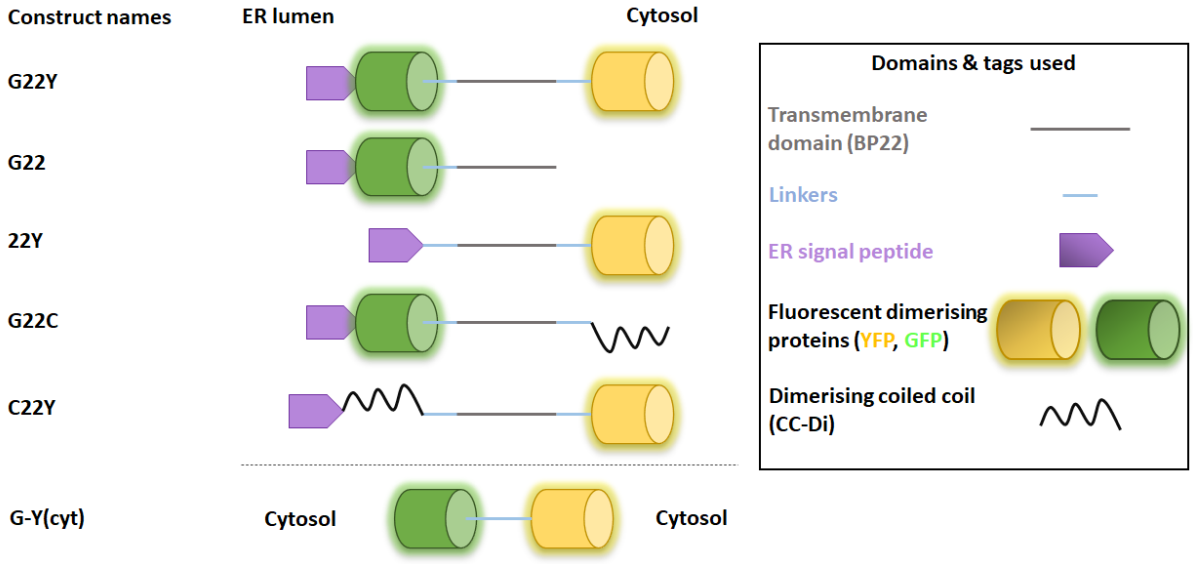


Figure 3.1: Schematic of the compartment forming polyprotein variants used. Five new polyprotein versions were generated from G22Y to study the effects of protein structure on compartment formation. The top five constructs contain BP22 (a transmembrane domain; in grey) and an ER signal peptide (purple), which targets the polyproteins to the ER membrane. These modules are removed in G-Y(cyt), yielding a cytosolic protein. The dimerising proteins used were GFP (green), YFP (yellow) and the synthetic coiled-coil CC-Di (black).

3.2.2 Initial characterisation of the six compartment forming constructs

All five G22Y-derived constructs (Figure 3.1) were successfully cloned (confirmed by Sanger sequencing of plasmids isolated from successfully transformed colonies) and all six constructs were transiently expressed in mature tobacco leaves using *A. tumefaciens*. Successful expression of all constructs was confirmed by visualisation of the fluorescent proteins using confocal microscopy and a variety of compartment morphologies were observed (Figure 3.2). G22Y, C22Y and 22Y constructs usually (in more than 90% of the transformed cells) formed a single large compartment (up to 25 µm for G22Y, up to 50 µm for 22Y and C22Y) in each transformed cell (Figure 3.2A-C). G22Y and C22Y formed more

compact, globular structures, with the former having a rougher edges and holes in the structure, which was not apparent for C22Y. In contrast, 22Y formed long, thick strands. G22 and G22C constructs formed a large number of smaller spherical compartments (Figure 3.2D-E). The removal of the transmembrane domain and the ER-localisation in G-Y(cyt) caused it to accumulate in the cytosol without forming any larger structures (Figure 3.2F). This suggests that association of the oligomerising protein with the ER membrane is essential for the formation or persistence of the compartments.

G22C constructs were co-expressed with a plasmid encoding the peroxisomal marker protein YFP-PEX2 and the spatially overlapping GFP and YFP signals suggested that the G22C proteins are mistargeted and enter the peroxisomes instead of forming novel structures (Figure 3.3). G22 showed a similar phenotype, but YFP-PEX2 co-expression was not successful due to low expression levels. Co-expression of the two constructs with the ER-lumenal marker RFP-HDEL showed no overlap (Figure 3.3), and these constructs were not investigated further.

To confirm that the compartments were derived from the ER, and to determine if they have a membrane and a lumen, G22Y, C22Y and 22Y were co-expressed with the ER markers RFP-HDEL (lumenal) and TAR2-RFP (membrane-bound) (Figure 3.4). All three constructs showed strong overlap with both ER markers, in line with the initial hypothesis that these compartments inherit their membranes and lumen from the ER. Analysis of the subcellular localisation of the RFP signal in cells co-expressing C22Y and TAR2-RFP showed that only an average of 9.2% of RFP signal was in compartments, suggesting that most of the ER membrane remains in its original state, outside of the compartment.

To establish the process of G22Y compartment formation, transformed leaf epidermal cells were imaged every 12 h until 168 h post-infiltration (Figure 3.5). Cells started to show fluorescent signals 36 h after infiltration, initially with a number of smaller structures (up to 5 µm in diameter) forming that then coalesced into one (or in some rare cases two) large compartments. While there was heterogeneity at the single cell level on the timing of this transformation (possibly due to varying expression levels and exact timing of *A. tumefaciens* mediated transformation), larger compartments started to appear around 48-60 h after infiltration. These large compartments remained stable for a long time, as cells expressing the fluorescent proteins remained viable until leaf senescence (up to 6 weeks in this study) (Figure 3.5), and normal cell behaviour (i.e., cytoplasmic streaming identical to untransformed cells) was observed until leaf senescence. Other compartment types also showed stability for several weeks (Figure S2), and transformed leaves had no negative macroscopic phenotype (Figure S3).

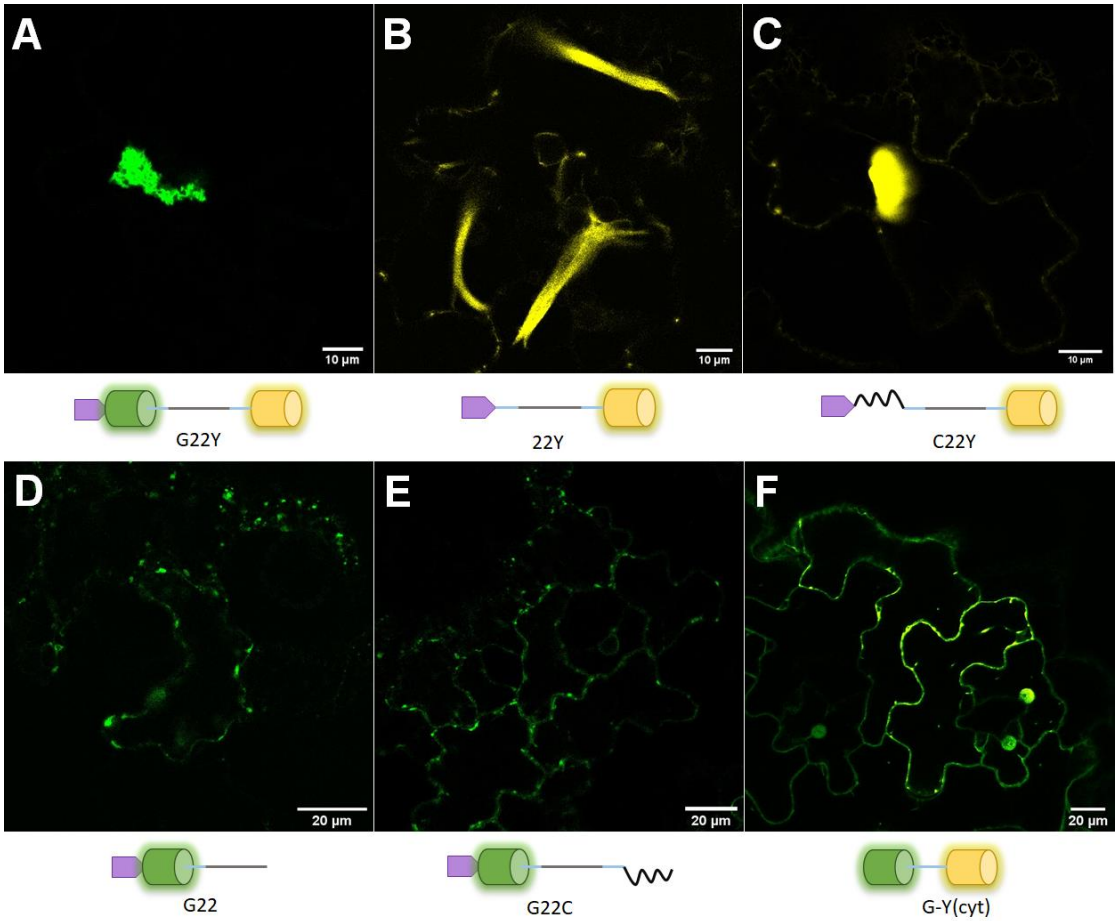


Figure 3.2: Confocal microscopy of the G22Y family constructs 5 days after transformation of mature tobacco leaves. Green: GFP. Yellow: YFP. The schematic of the corresponding construct from Figure 3.1 is shown under each image. A: A single G22Y compartment (only GFP is visualised) of about 25 µm. Note the uneven edges of the compartment. B: Elongated 22Y compartments. C: C22Y compartment, with more uniform and consistent labelling by the YFP signal. D&E: G22 (D) and G22C (E) constructs do not coalesce into a single large compartment, instead assembling into several smaller ones. F: G-Y(cyt) constructs do not form cytosolic compartments but result in typical cytosolic and nuclear labelling patterns. Scale bars: 10 µm in A-C, 20 µm in D-F.

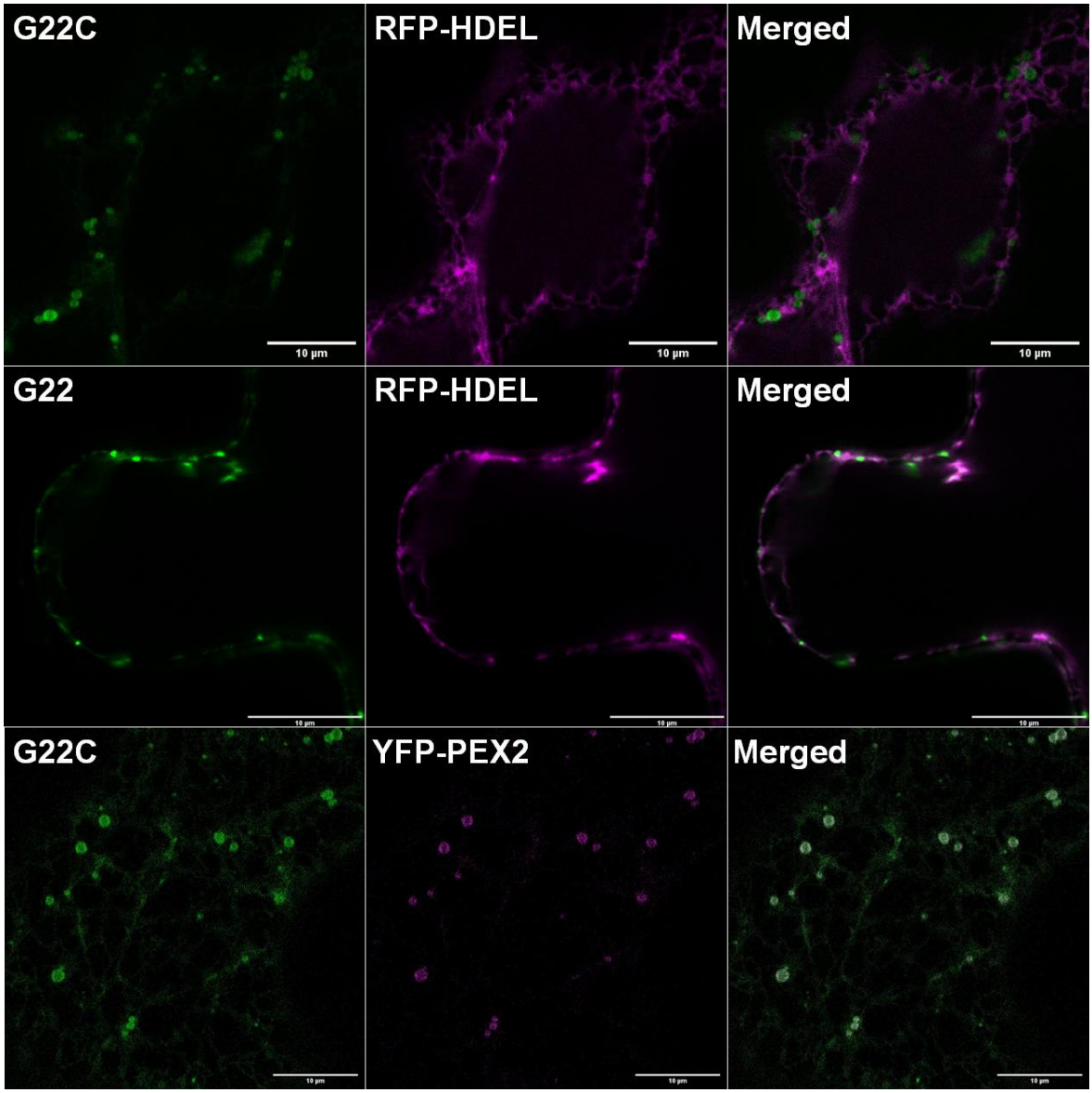
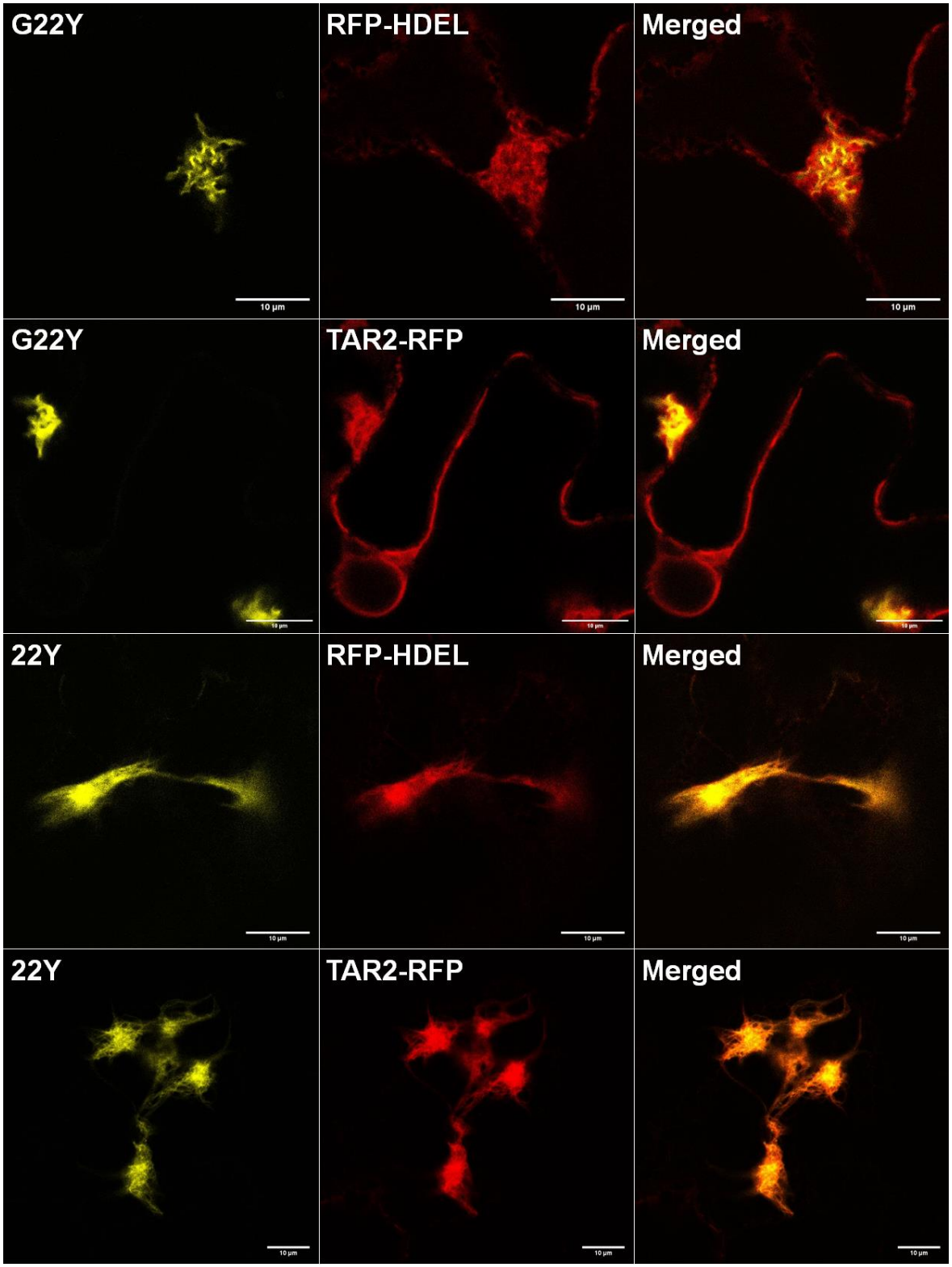


Figure 3.3: Co-expression G22C and G22 with fluorescent subcellular markers in mature tobacco leaves. YFP-PEX2 is a peroxisomal marker and RFP-HDEL is an ER lumenal marker. YFP-PEX2 highlights the small compartments produced by G22C, but RFP-HDEL does not for either G22 or G22C. G22 – YFP-PEX2 co-expressions were unsuccessful due to low expression. Images were captured using fluorescent confocal microscopy 5 days after transformation by *A. tumefaciens*. Scale bars: 10 µm


G22Y
RFP-HDEL
Merged
G22Y
TAR2-RFP
Merged
22Y
RFP-HDEL
Merged
22Y
TAR2-RFP
Merged

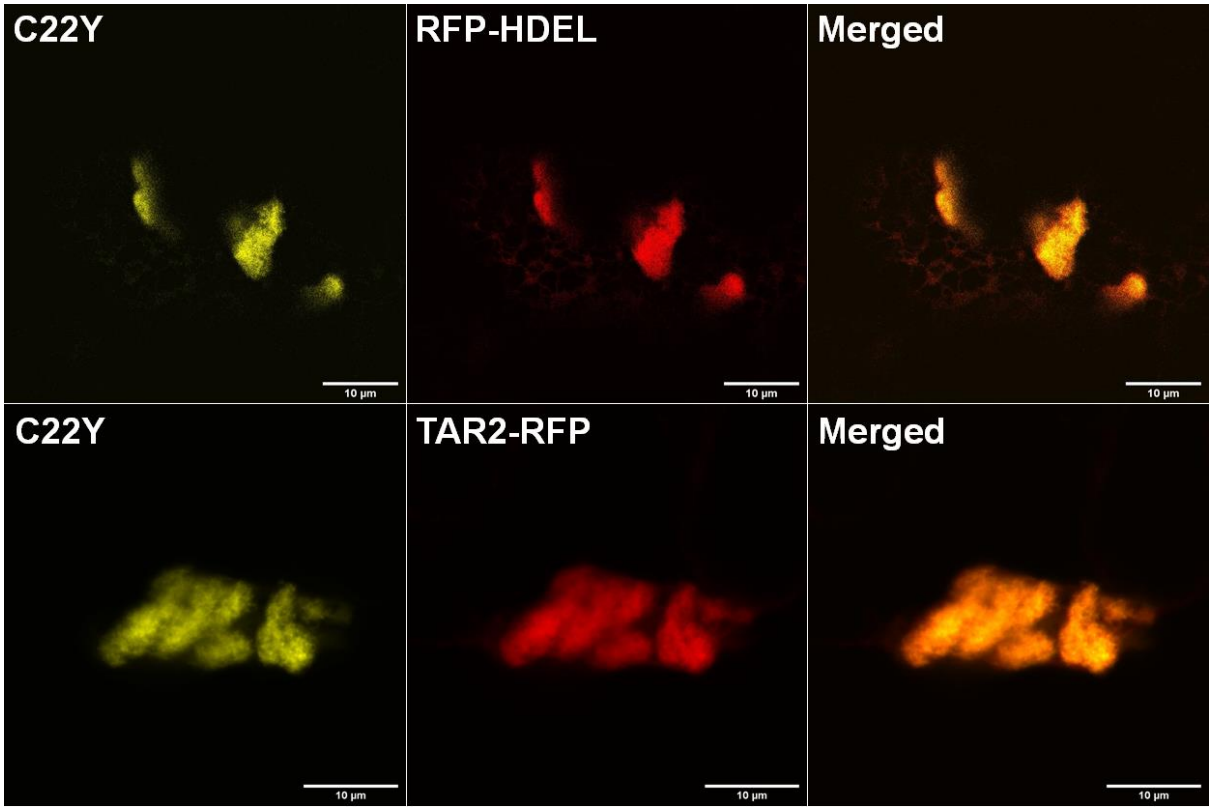
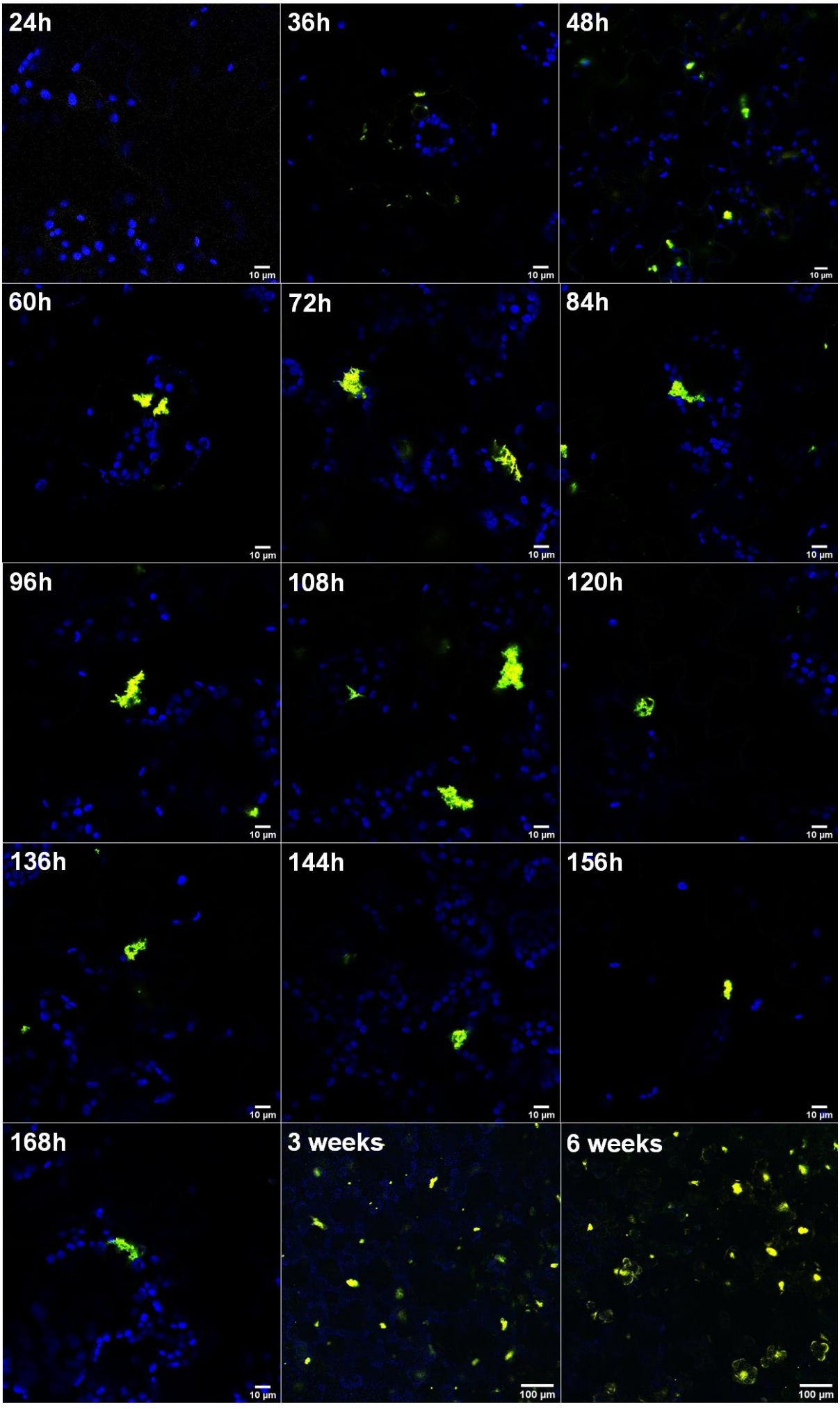


Figure 3.4 (above and previous page): Co-expression of ER markers with the synthetic compartments, to determine if the compartments inherit their lumen and membrane from the ER. TAR2-RFP is an ER membrane marker. Both the ER lumen and ER membrane markers (red) overlap the fluorescent signal from the three synthetic compartments (yellow). The images were visualised using fluorescent confocal microscopy on mature tobacco leaves, 7 days after transformation by *A. tumefaciens*. Scale bars: 10 µm.

Figure 3.5 (below): Expression time-course of G22Y from 24 h to 6 weeks in mature tobacco leaves. GFP signal is shown as green, YFP signal as yellow and chloroplast autofluorescence as blue. GFP and YFP signals remain completely co-localised and typically one large compartment per cell was detected from 48-60 h post-transformation. No obvious signs of cellular toxicity were detected throughout the time series, which ran until leaf senescence. No expression of G22Y was detected at 24 h or before. The images were visualised using fluorescent confocal microscopy. Scale bars: 24 h - 168 h: 10 µm, 3 weeks & 6 weeks: 100 µm.


24h
36h
48h
60h
72h
84h
96h
108h
120h
136h
144h
156h
168h
3 weeks
6 weeks
10 µm
100 µm

3.2.3 High-resolution analysis of compartment structures

The three constructs producing large compartments (G22Y, 22Y and C22Y) were visualised using higher resolution airy-scanning confocal microscopy (Figure 3.6). All three formed large, but morphologically distinct structures: G22Y compartments showed helical structuring (Figure 3.6A & Supplementary Figure S4A-B), while 22Y formed long fibrous strands (Figure 3.6B) and C22Y presented a more homogenous YFP signal which could not be further resolved using airy-scanning confocal microscopy (Figure 3.6C). G22Y helices were measured to determine the turn length of a helix. Helix turn length was found to be variable, with values ranging between approximately 0.75 - 1.6 µm (Supplementary Figure S4C).

To further characterise the internal structures of these large compartments TEM was used. Unfortunately, due to low expression, TEM of G22Y was unsuccessful. TEM images of 22Y (Figures 3.7A-B) showed details of the fibrous structures (Figures 3.7B) which appeared to be an ER cylinder with a tube of cytosol trapped in the centre (Figure 3.7C). C22Y compartments presented less clear patterns of organisation, with more complex and densely packed membrane structures observed (Figure 3.7D&E). The density of membranes explains why the structure of C22Y compartments could not be resolved by high-resolution airy-scanning confocal microscopy, as the ER membranes would be in too close proximity to be individually resolved by this method. However, 22Y strands are more separated, allowing clear resolution using confocal microscopy as well.

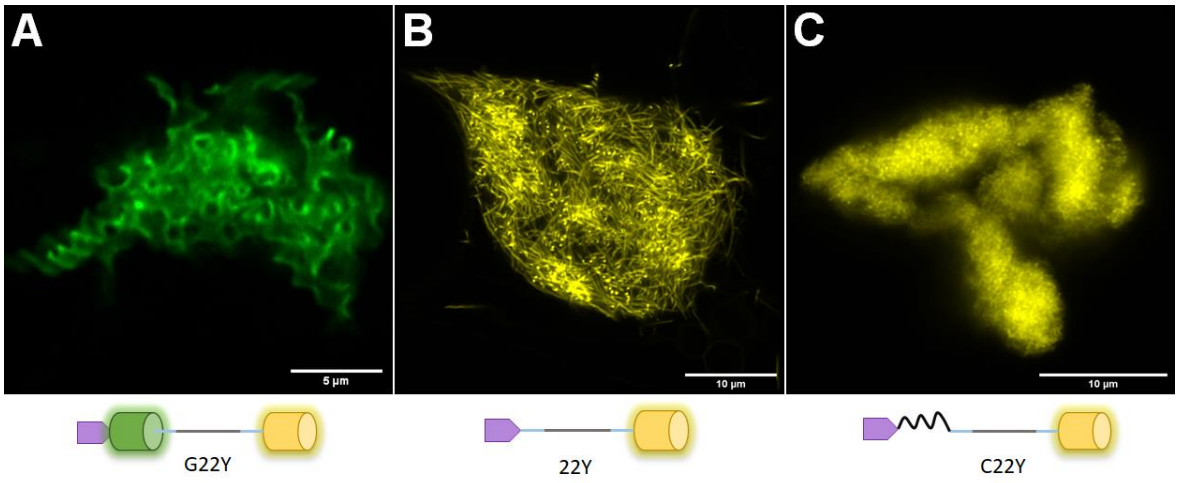


Figure 3.6: High-resolution airy-scanning confocal microscopy images of the G22Y family of constructs 5 days after transformation of mature tobacco leaves. Green: GFP. Yellow: YFP. The schematic of the corresponding construct from Figure 3.1 is depicted under each image. A: A G22Y compartment of about 20 µm visualised using the GFP signal. Note the angle of the edges and circular holes suggesting some form of helical structure. B: A 22Y compartment with fibrous structure. C: A C22Y compartment with more consistent and complete coverage of the compartment by the YFP signal. Scale bars: 5 µm in A; 10 µm in B&C.

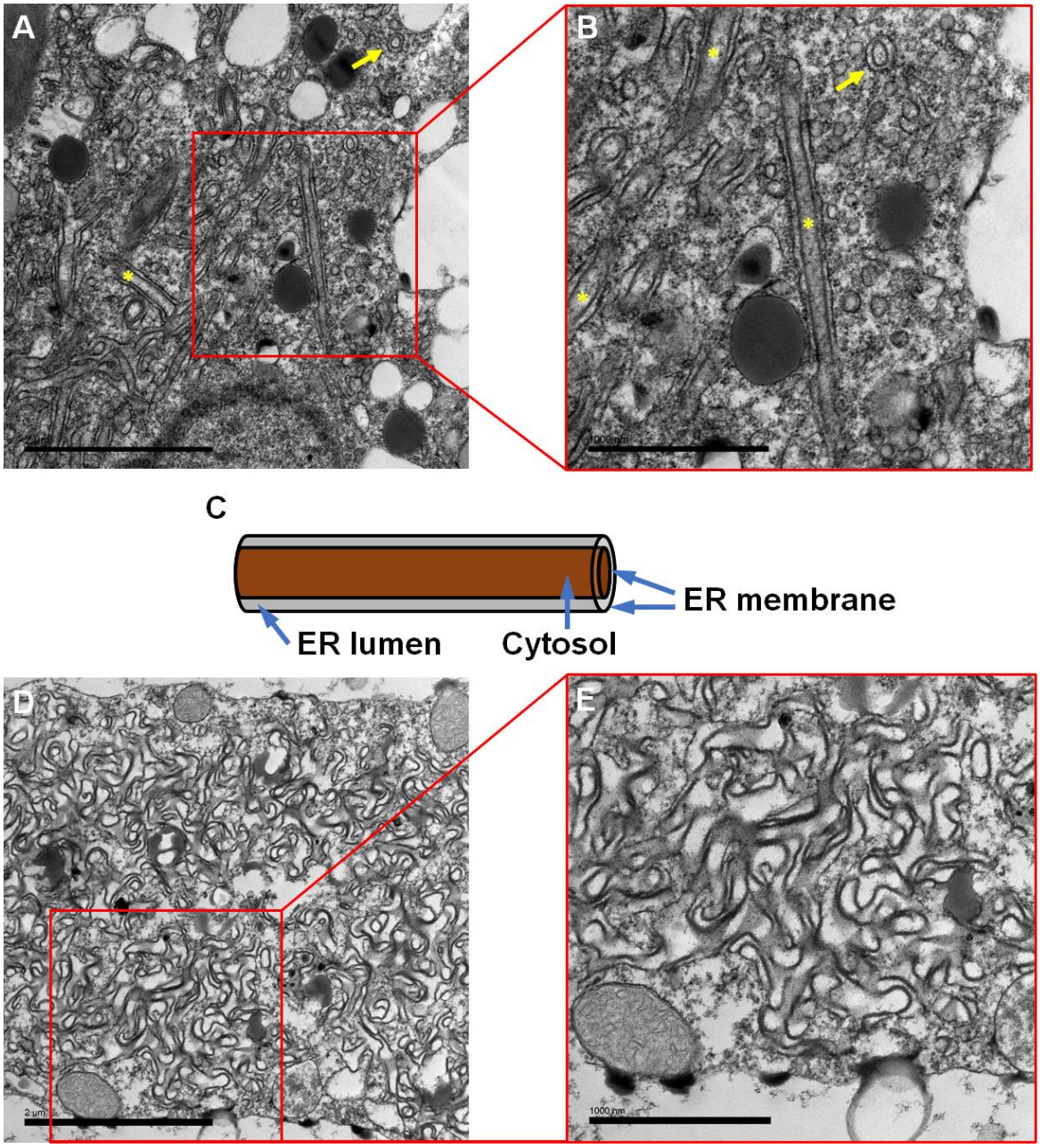


Figure 3.7: TEM of 22Y and C22Y expressing tobacco epidermis cells. A&B: 22Y structures often presented as two closely aligned (average width of 50nm) parallel membrane pairs forming a long four-membrane structure with a larger gap (average width of 200nm) between the membrane pairs (yellow asterisks). This is probably an ER cylinder, and some images showed a cross-section of these from a different angle (yellow arrows). C: The schematic of the proposed ER cylinder structure (ER lumen in grey, ER membrane in black), surrounding the trapped cytosol in the centre (in brown). D&E: The C22Y structures are more complex, with membranes often forming pairs of parallel, highly curved and very densely packed networks. Scale bars: A&C: 2 µm, B&D: 1 µm. Due to time-constraints, this experiment was only carried out once, and should be repeated to confirm the results. Images were taken by Errin Johnson (Dunn School of Pathology, University of Oxford).

3.2.4 Compartment topology and diffusion dynamics

To investigate if the compartments were still connected to the ER network, the lateral diffusion and exchange of proteins were assessed using FRAP. The compartment forming constructs were co-expressed with ER lumenal RFP-HDEL and then photobleached. When the compartments were completely or partially photobleached, there was rapid recovery of the RFP-HDEL fluorescent signal in the compartments, confirming that the compartments were still connected to the ER (Figure 3.8). However, the rate of recovery for the lumenal marker was significantly retarded compared to the rest of the ER when entire compartments were photobleached (Figure 3.9A). The complementary FRAP of the compartment forming proteins (G22Y, C22Y, 22Y) embedded in the membrane showed no recovery, suggesting that these are immobilised due to the large number of connections in the compartment (Figure 3.8).

Another unresolved question was whether these compartments have cytosol trapped within their complex membrane structures. To investigate this, the compartments were co-expressed with the cytosolic fluorescent marker Peredox-mCherry, which was observed in both the cytosol and the compartments, confirming the presence of cytosol within the compartments (Figure 3.10). C22Y compartments showed slightly lower levels of the cytosolic probe inside the compartment compared to the bulk cytoplasm, suggesting that while some cytosol is trapped inside the compartment, most of the space is occupied by ER-derived membranes and lumen (Figure 3.10), which fits with the TEM results (Figure 3.7). FRAP of the Peredox-mCherry signal inside the compartments showed a similar retardation in the rate of recovery for the cytosol inside the compartments as for the ER lumen (Figure 3.9B). This confirmed that the cytosol trapped between the complex internal

membrane structures of the compartment was still connected to the bulk cytoplasm, but the rate of exchange was reduced.

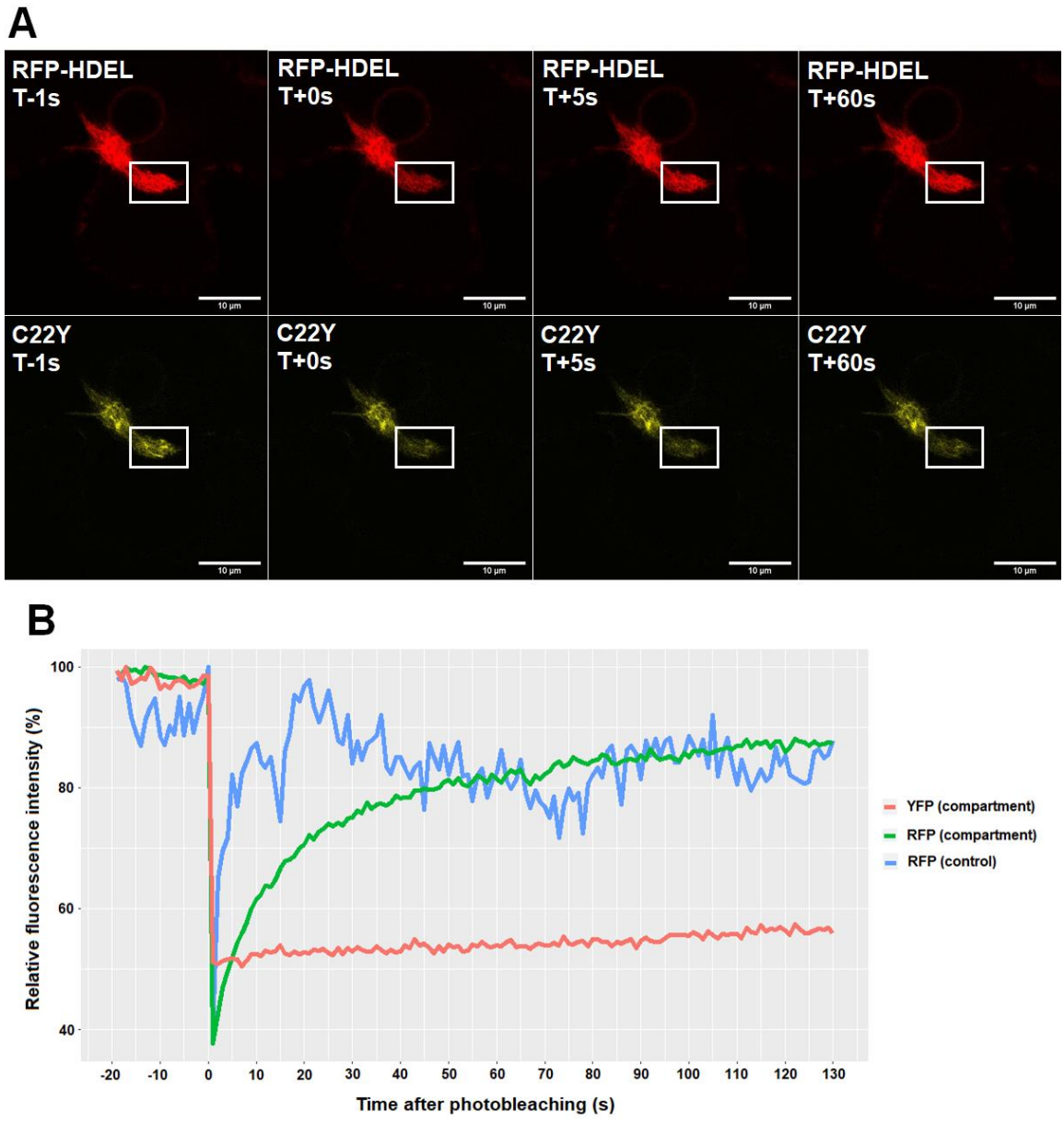


Figure 3.8: A representative example of FRAP of both the lumenal ER marker RFP-HDEL within and the YFP signal of the synthetic compartment. A: Confocal microscopy image series of FRAP. White box: region of photobleaching. T-1s: image taken immediately before FRAP. T+0s: image taken immediately after bleaching. T+5s: image taken 5 sec after bleaching. T+60s: image taken 60 sec after bleaching. Top row: RFP-HDEL. Bottom row: YFP signal from C22Y. The ER lumenal marker RFP-HDEL recovers quickly, with obvious increase in fluorescence even after 5 sec, while the compartment-forming C22Y does not recover even after 60 sec. Images were collected using fluorescent confocal microscopy on mature tobacco leaves, 5 d after transformation by *A.*

tumefaciens. Scale bars: 10 µm. B: Graphical representation of typical fluorescent recovery patterns after photobleaching. YFP signal (red) from the compartment forming polyproteins show very little recovery when compared to the RFP-HDEL signal. The ER luminal marker recovers slower to the compartment (green) than the rest of the ER (blue).

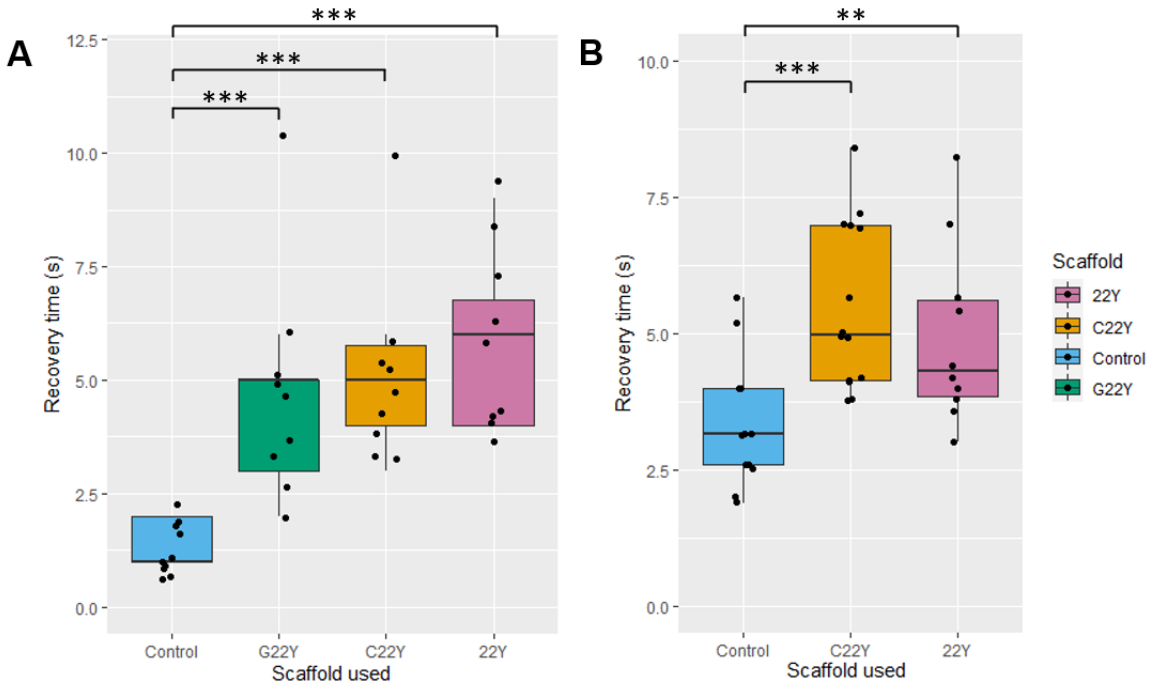


Figure 3.9: Retarded lumenal and cytosolic recovery rate in compartments after FRAP. A: The 50% normalised recovery time of RFP-HDEL inside the compartments after FRAP was significantly longer ($p < 0.001$) for all three constructs compared to the control (RFP-HDEL only). Comparison of compartment types showed no statistical difference for either pairing ($p > 0.4$ in all cases). $n = 9$ for all constructs (each datapoint is depicted with a black dot). B: Cytosolic recovery rates of Peredox-mCherry were also significantly retarded ($p < 0.01$ for 22Y and $p < 0.001$ for C22Y) compared to the control. The two compartment types did not differ significantly ($p > 0.15$). Asterisks depict significance levels: * for $p < 0.05$, ** for $p < 0.01$ and *** for $p < 0.001$. $n = 13$ for all constructs. The boxes represent the interquartile range, the horizontal line in the box shows the median, and the whiskers the minimum and maximum values (excluding outliers). Due to its low expression, it was not possible to quantify FRAP for G22Y – Peredox-mCherry co-expressions.

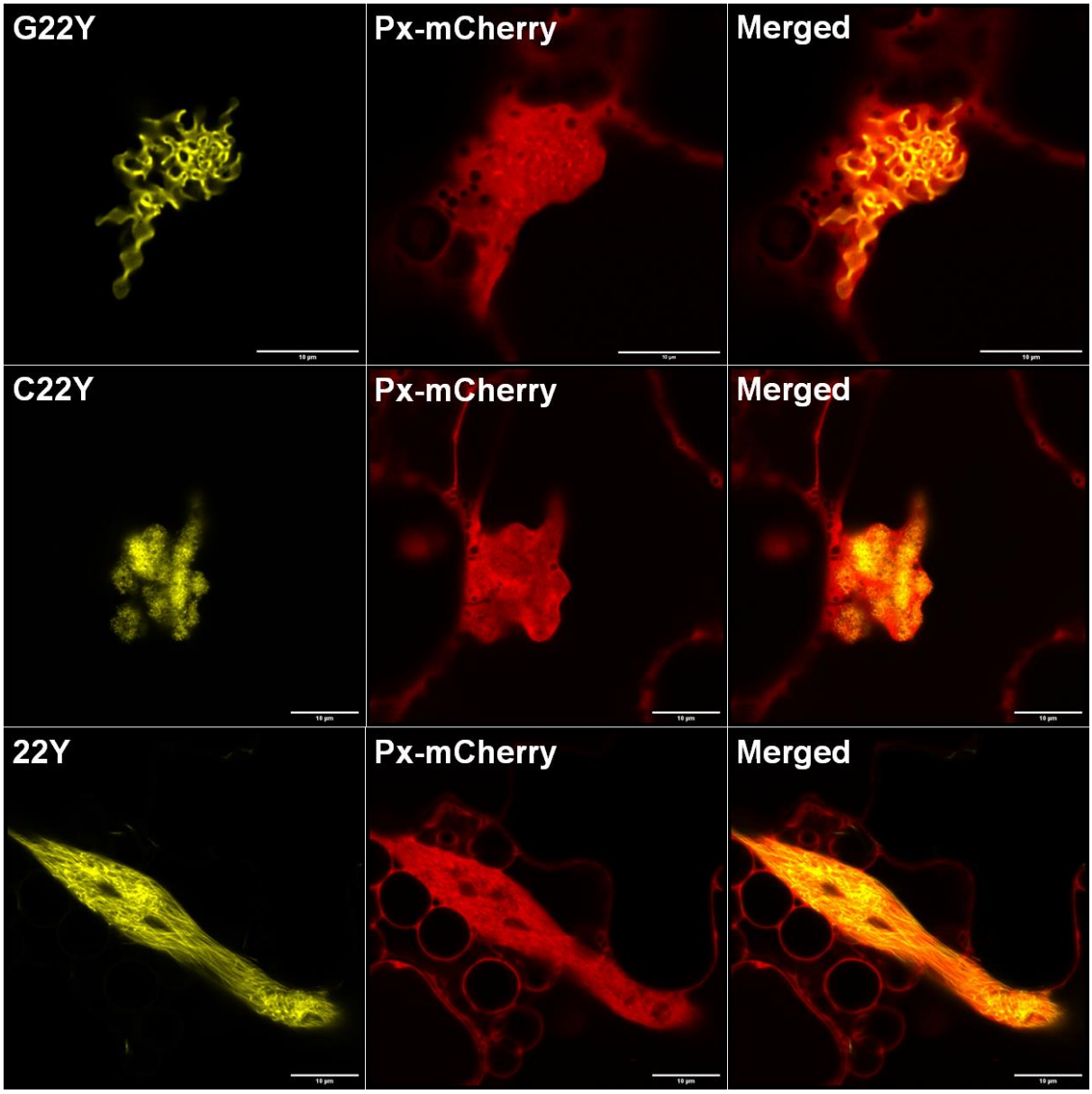


Figure 3.10: Visualising the cytosol trapped within the membranes of the compartments. Cytosolic Peredox-mCherry (Px-mCherry, red) was co-expressed with the synthetic compartments, visualised using YFP fluorescence (yellow). Note C22Y images, where it is clear that the compartment areas have slightly lower mCherry signal than the surrounding cytosol, suggesting that C22Y compartments have less trapped cytosol, and probably ER lumen and membranes are more concentrated. The images were collected using high-resolution airy-scanning confocal microscopy on mature tobacco leaves, 7 days after transformation by *A. tumefaciens*. Scale bars: 10 µm.

Notably, cells with 22Y often form a number of large spheres in close proximity to the main compartment in the centre of the cell, with YFP signal present in the membrane of the spherical structures (Figure 3.11 & Supplementary Video SV1). Co-expression with the vacuolar membrane marker GFP-TIP1:1 showed that the surface of the spheres also contained this marker, however at a lower level than the rest of the vacuolar membrane (Figure 3.11), and that the vacuolar membrane curves around the spheres, suggesting that they are not inside the central vacuole. When co-expressed with fluorescent ER probes, TAR2-RFP was found in the membrane of the spheres (Figure 3.11) and RFP-HDEL was concentrated inside the spheres (Figure 3.11), suggesting that these structures are derived from the ER. FRAP using this RFP-HDEL showed that there was no connection between the lumen of individual spheres. The cytosolic probe Peredox-mCherry also highlighted the membrane of the spheres (Figure 3.10 & Figure 3.11). This could be a cytosolic 'halo' due to cytosol entering the space between the sphere and the vacuolar membrane.

Combined, these results suggest that these structures are either swollen ER bodies previously described[179], found now in fully developed leaf tissue or possibly smaller vacuoles produced due to ER stress.

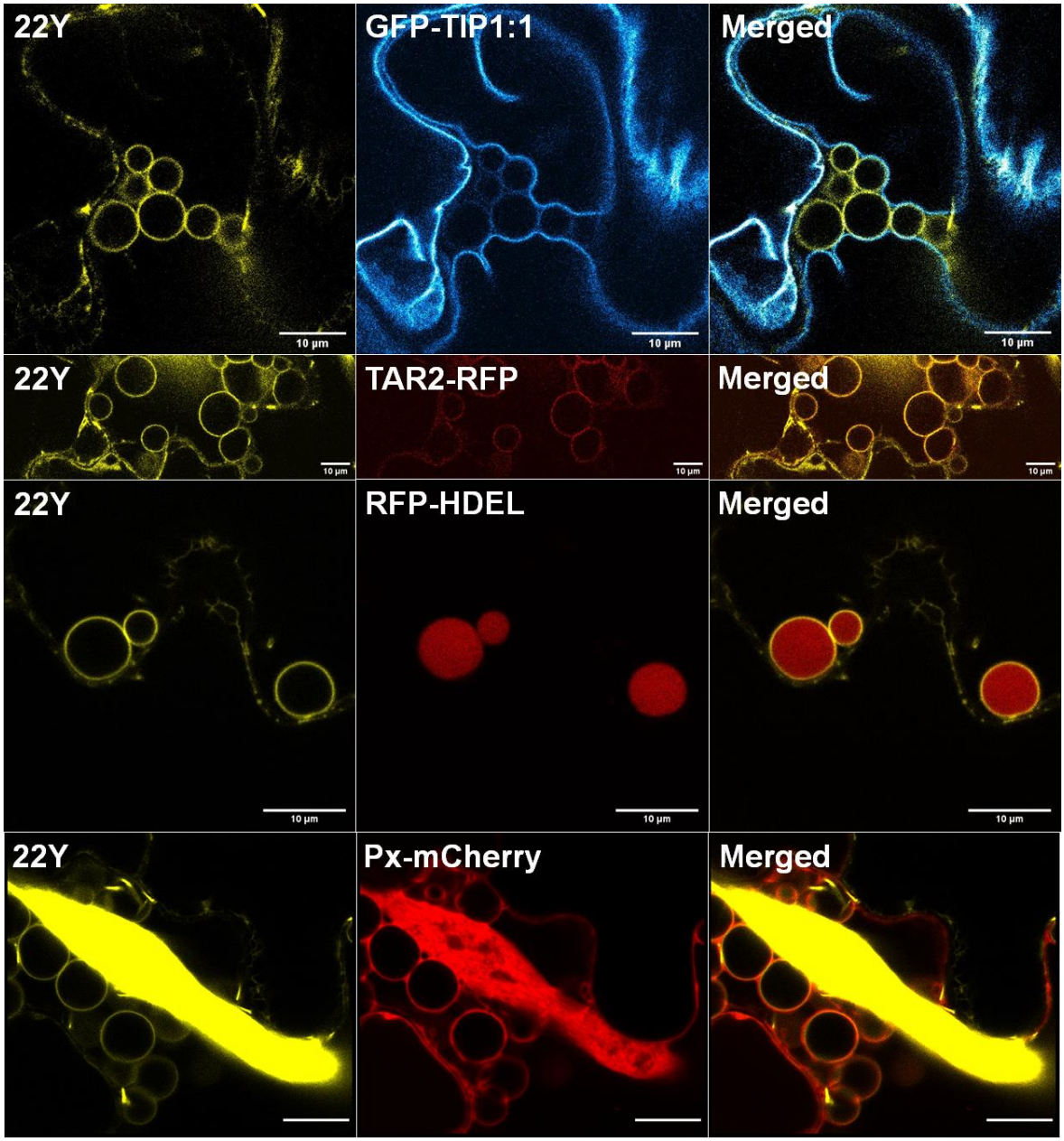


Figure 3.11: Co-expression of tonoplast and ER markers with 22Y spheres (yellow) in mature tobacco leaves. GFP-TIP1:1 (blue) is a tonoplast marker. GFP-TIP1:1 is present in the membrane of the spheres, but at lower quantity than in tonoplast membranes, which push against the outer membrane of the spheres. TAR2-RFP is in the membrane of the spheres, and RFP-HDEL accumulates inside the spheres instead of the ER. When 22Y fluorescence signal is increased in brightness on Figure 3.10, the spheres become visible. Scale bars: 10 µm. The images were visualised using fluorescent confocal microscopy, 7 days after transformation by *A. tumefaciens*.

3.2.5 Quantitative ER network architecture and dynamics

To assess if the introduction of the three types of compartments caused any disturbance to the rest of the ER, the ER (excluding the compartments) was visualised in mature tobacco leaves 7 days after transient transformation with the RFP-HDEL marker protein and one of the compartment types using confocal laser scanning microscopy. The marker highlighted that the non-compartmentalised ER shows no aberrant phenotypes on a visual inspection.

ER network architecture and dynamics were captured using short (20 s) time-series images that were quantitatively analysed using the AnalyzER software package[168]. 41 ER parameters were investigated, including topological, morphological and kinetic parameters for the cisternae, tubules and polygonal regions enclosed by the ER network (Table 3.1). The data was subjected to PCA, and the principal components were individually assessed for any statistically significant differences (using MANOVA and ANOVAs) between compartment-expressing cells and controls. The first two principal components (PC1 and PC2, which accounts for 32.0% and 18.4% of variance in the data, respectively) showed statistically significant differences between the groups ($p = 0.015$ for PC1 and $p = 0.009$ for PC2). However, when further investigated using the Tukey's HSD test, it became apparent that most of this variance was caused by differences between compartment types, and not between the wild-type and the compartments, which is in line with the visual inspection of the PCA plot (Figure 3.12).

When compared to the wild type (WT), only C22Y showed a statistically significant difference ($p = 0.041$) and only for PC2. To investigate how the ER had been disrupted in this case, the 17 ER parameters that contributed the most to PC2 were individually tested for any significant changes (Table 3.1), but none of them reached significance ($\alpha=0.05$). To

make sure that no parameter was excluded unduly by this, all 41 parameters were individually tested later using ANOVA and Tukey's HSD test, but the latter showed no significant differences between WT and either compartment type (Supplementary Table S1). This suggests that individually their effect size was very small, and that no individual parameter of the ER had been statistically significantly altered. This, in combination with the finding that cells with compartments remain viable until leaf senescence, suggest that the ER can readily compensate for any disruption caused by the introduction of the novel compartment, and its effect on the cell is minimal.

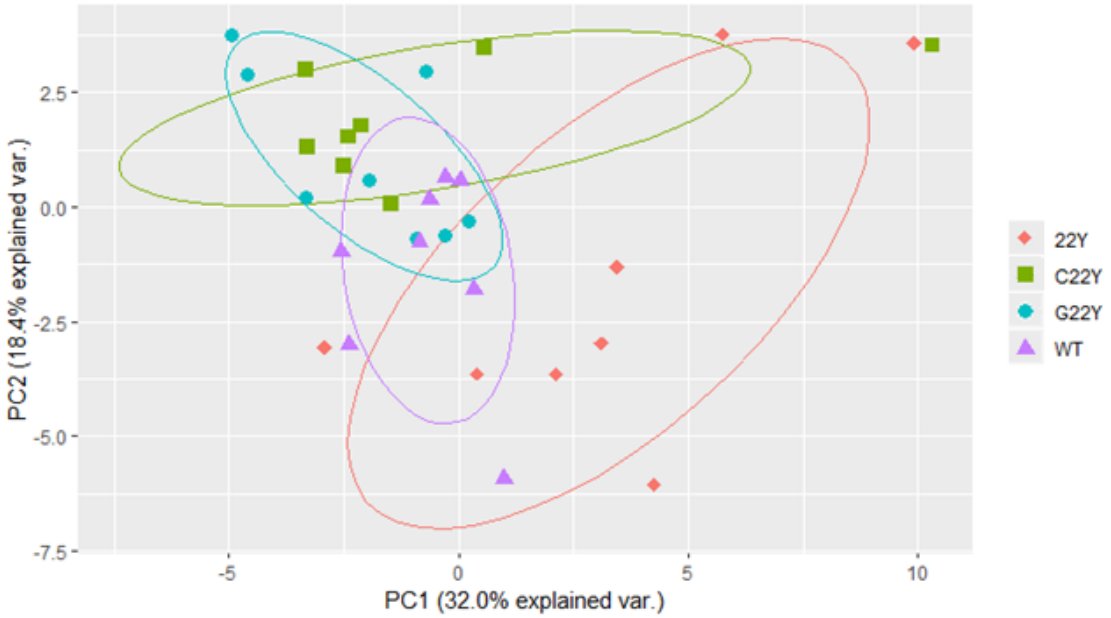


Figure 3.12: Principal component analysis of the ER network dynamics. The two primary principal components are shown here, which explain the greatest proportion of the variance in the data. All groups overlap, but there is less overlap between different compartment types, especially 22Y and the other two (C22Y and G22Y). This causes the MANOVA to be statistically significant, but when only comparing compartment types to the WT (and not to each other), most of this variance is lost. n = 8 for all categories

Table 3.1 (below): Parameters used in the AnalyzER software package to investigate network dynamics. Descriptions are from the AnalyzER manual (version 1.1). Parameters marked with * refer to the 17 parameters that contributed the most variance to PC2 and were further investigated for significant differences between groups.

Parameter type	Parameter name	Description
Tubule edges	Length	The total length of tubules (l, µm)
Tubule edges	Volume*	The total cross-sectional volume of tubules ($v = a * l$, μm^3)
Tubule edges	Resistance	The predicted resistance to flow assuming Poiseuille flow (l/r^4, μm^{-3})
Tubule edges	Tortuosity	The Euclidean distance between the nodes divided by the total length of the edge
Tubule edges	Width_center*	The estimated width excluding overlap regions at the node
Tubule edges	Speed_local	The scalar sum of the speeds calculated for each pixel ($\mu m\ s^{-1}$)
Tubule edges	Speed_max	The maximum speed for any pixel ($\mu m\ s^{-1}$)
Tubule edges	Speed_global	The vector sum of speeds for each pixel ($\mu m\ s^{-1}$)
Tubule edges	Flow_coherence	The ratio of the local speed (scalar sum) to the global speed (vector sum)
Tubule edges	Persistency	The mean period of time that each pixel forms part of a tubule (s)
Tubule nodes	node_Omin_Omaj	Branch angles between the three main incident tubules, determined from a linear segment from the node to the midpoint of each tubule
Tubule nodes	node_Omid_Omaj	
Tubule nodes	node_Omin_Omid	
Tubule nodes	node_Speed	Sum of speed at nodes ($\mu m\ s^{-1}$)
Tubule nodes	node_Persistency	The mean period of time that each pixel forms part of a node (s)
Cisternae	cisternal_Node_degree*	The number of connecting tubules incident on the cisterna
Cisternae	cisternal_Contrast_all*	A measure of the intensity contrast between pixel i and its neighbour j. Values range from 0 to $(nbins1)^2$. Here, results are normalised to $(nbins-1)^2$ to fall between [0 1]. An idealised cisternal sheet would have a contrast of zero.
Cisternae	cisternal_Correlation_all*	A measure of how correlated a pixel is to its neighbour. where µi and µj are the weighted mean intensities, and σi and σj are the standard deviations of the GLCM distributions. Values range from -1 (un-correlated) to 1 (fully correlated). An idealised cisternal sheet would have a value of 1.
Cisternae	cisternal_Energy_all	Gives the sum of squared elements in the GLCM. Values range from 0 to 1. An idealised cisternal sheet would have an energy of 0.
Cisternae	cisternal_Homogeneity_all*	Measures the closeness of the distribution of elements in the GLCM to the diagonal. Values range from 0 to 1. An idealised cisternal sheet would have a value of 1
Cisternae	cisternal_Speed_local	The scalar sum of local speeds at every pixel ($\mu m\ s^{-1}$)
Cisternae	cisternal_Speed_global*	The vector sum of global speed ($\mu m\ s^{-1}$)
Cisternae	cisternal_Flow_coherence*	The ratio of the local speed (scalar sum) to the global speed (vector sum
Cisternae	cisternal_MaxPersistency*	The maximum persistency of pixels in each cisterna (s)
Cisternae	cisternal_MeanPersistency*	The average persistency of pixels in each cisterna (s)
Cisternae	cisternal_MaxDistance*	The maximum distance of any pixel to the edge of the cisterna (µm)
Cisternae	cisternal_MeanDistance*	The average distance of any pixel to the edge of the cisterna (µm)
Cisternae	cisternal_Area*	The area of each cisterna
Cisternae	cisternal_Solidity	The proportion of the pixels in the convex hull that are also in the cisterna
Cisternae	cisternal_Elongation	The ratio of the major axis to the the minor axis
Cisternae	cisternal_Roughness	The ratio of the $perimeter^2$ to the area
Cisternae	cisternal_circularity	The ratio of the radius determined from the area to the radius determined from the perimeter
Polygons	polygon_Area	The area of each polygonal region with the ER tubules thinned to a single-pixel wide skeleton
Polygons	polygon_Solidity*	The proportion of the pixels in the convex hull that are also in the region
Polygons	polygon_Circularity	The ratio of the radius determined from the area to the radius determined from the perimeter
Polygons	polygon_Roughness*	The ratio of the $perimeter^2$ to the area
Polygons	polygon_Elongation*	The ratio of the major axis to the the minor axis
Polygons	polygon_MaxDistance	The furthest distance within the region to the ER network
Polygons	polygon_MeanDistance	The average distance within the region to the ER network
Graph	Geff	The global efficiency of the network
Graph	Alpha*	The alpha coefficient or meshedness

3.3 Discussion

3.3.1 The effect of the orientation of the dimerization domains on the morphology of the compartments

A notable observation for the six compartment-forming construct types investigated here was the wide variety of compartment morphologies. G22 and G22C were mistargeted to peroxisomes, G-Y(cyt) did not form a compartment, while every other construct generated its own unique compartment type. This shows that the nature of the dimerising protein domains (both the cytosolic and the lumenal facing) clearly affect the compartment structure, and that a transmembrane domain is essential for compartment formation.

A working hypothesis to explain the connection between the polyproteins and compartments ultrastructure is shown in Figure 3.13. The relative position of the transmembrane domain facing ends of the dimerising protein pairs determine the predicted membrane complexity (Figure 3.13). For the protein constructs investigated here, the relative orientation of the C-terminal ends of the lumenal dimerising protein pairs, and the N-terminal ends of the cytosolic dimerising protein pairs were critical.

As GFP and YFP form antiparallel dimers, the C-terminus of their dimers face in opposite directions. This likely allows the two attached transmembrane domains to flexibly bind either to the same membranes or an adjacent membrane (Figure 3.13). When only a single dimerising protein was present (22Y), this can bridge two adjacent membranes in close proximity (Figure 3.13), but when dimerising domains are present on both sides of the transmembrane domain (G22Y) a more complex membrane system can be assembled (Figure 3.13). The observed helical structure is likely due to the introduction of a small, repeating angle in the membrane curvature (probably due to the construct's bindings on

either side of the membrane). Note that the flexibility of GFP in binding to different membranes was not originally assumed by Moore and Samalova (unpublished), hence the different schematic representation of the G22Y oligomer here compared to Figure 1.4. However, there is no reason to believe this flexibility is prohibited by the structures of the transmembrane domain and the linkers. An introduction of a helix could also have significant impact on the density of the membrane structure, since the repeating pattern would cause the internal ultrastructure of the compartment to have regularly interspaced membrane-free areas, which would be filled with cytosol.

The coiled-coil domain, which replaces the lumenal GFP in C22Y, has a different structure: it binds in parallel, causing the C-termini of the dimers to be oriented in the same direction. This restricts the possible membrane binding sites for the transmembrane domain, and due to the short linker will force the dimers to bind to the same membrane, which in turn restricts the possible membrane configurations (Figure 3.13). However, this removes the restriction imposed by the helical membrane structure on the compartment, and a less organised, but more densely packed membranous structure is possible (when compared to 22Y), as observed by TEM (Figure 3.7). This hypothesis is supported by the earlier finding by Moore *et al.*, that replacing the C-terminal GFP of G22Y with an mRFP, gave a phenotype identical to C22Y[156]. However, it was later found that mRFP forms parallel binding dimers[180]. These dimers have C-termini that also face the same direction, like the coiled-coil domains, but unlike GFP or YFP (Figure 3.13).

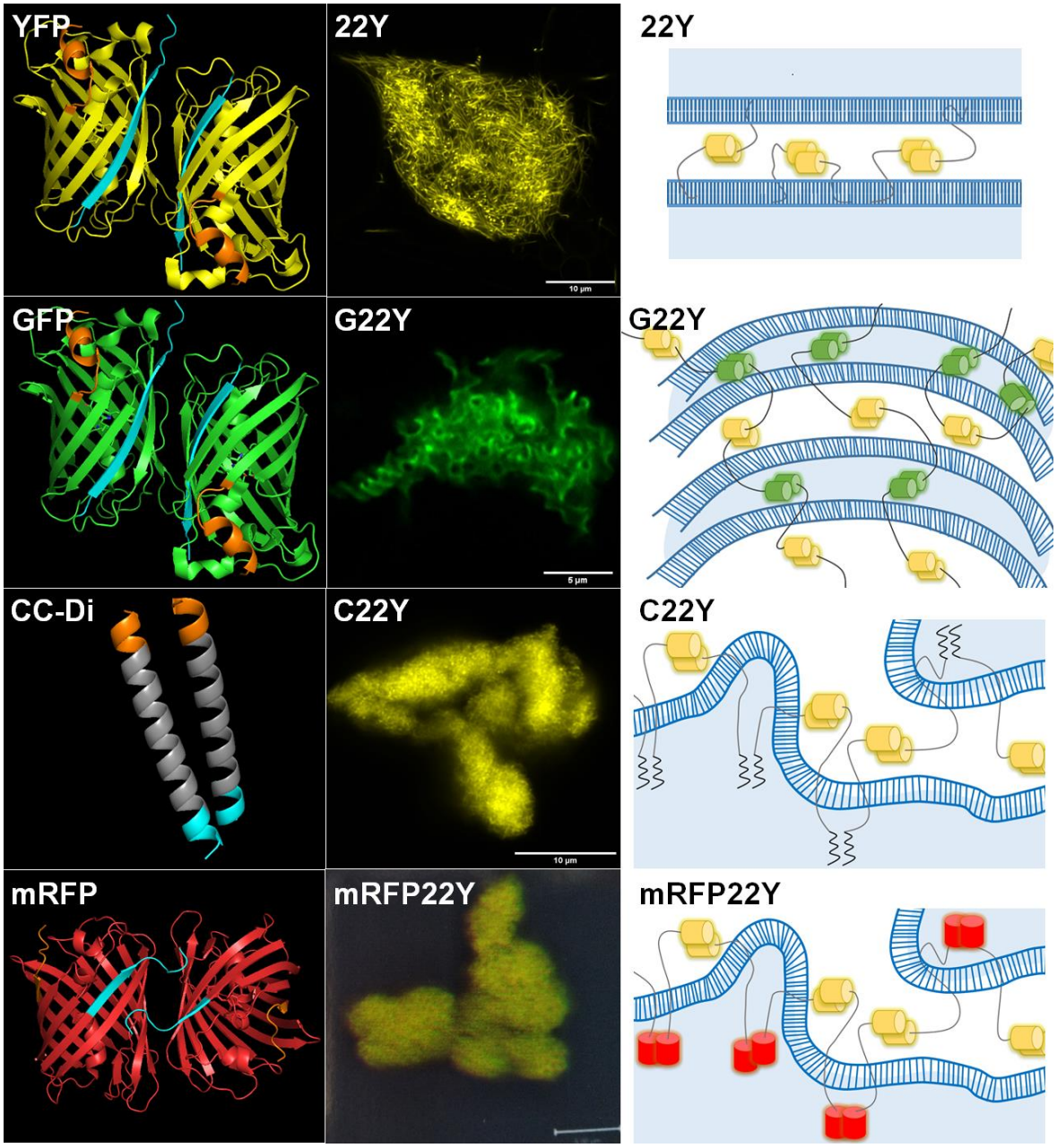


Figure 3.13: Relative C-terminal positioning of dimerising proteins affects the structure of the synthetic compartment. First column: molecular structure of the dimerising proteins with the C-terminus highlighted in cyan and N-terminus highlighted in orange. Note that the C-termini in YFP and GFP dimers are distal facing (antiparallel dimerization), while for CC-Di and mRFP these are proximal facing (parallel dimerization). This correlates with the changed structure of the compartments in the second column (confocal microscopy images). The third column show comparative diagrams of the proposed membrane interactions. ER lumen is shaded light blue, cytosol is white. The schematics show how distal facing C-termini of dimers allow interactions between multiple membranes, which increases the complexity of the compartment when more dimerising proteins are introduced in G22Y, causing the formation of helical structures. When GFP

is replaced by a dimerising protein with a proximal facing C-terminus, the proximal facing C-termini restrict the dimer to the same membrane and a less structured, but more densely packed membrane structure is observed (as seen by TEM results). Confocal microscopy image of mRFP22Y is from White (2006). Scale bars: 22Y & C22Y: 10 μm, G22Y & mRFP22Y: 5 μm.

The same principles also explain the differences in the observed compartment morphologies for polyproteins where the cytosolic dimerising domains were replaced. When the cytosolic dimerising YFP was removed in G22, no compartment was formed highlighting the importance of the cytosolic dimerising domains. The cytosolic YFP in G22Y, C22Y and 22Y dimerises in an antiparallel orientation, with N-termini facing in opposite directions, allowing connections to form between membranes and in turn, the formation of membrane stacks. However, in G22C when YFP was replaced with coiled-coils which can only form dimers within the same membrane sheet, we get a structure similar to G22, as no membrane stacks can form in either, and G22C proteins are mistargeted to the peroxisomes (Figure 3.3). Furthermore, when previously the cytosolic YFP was replaced mRFP, which can only form dimers within the same membrane, similar to the coiled-coil domains, the same morphology was observed as for G22 and G22C[156].

Due to the modular nature of the polyproteins, this hypothesis is fully testable, for example by the introduction of antiparallel coiled-coil domains to replace GFP or YFP, where a similar compartment would be expected to form. More work on this topic is described in detail in Chapter 4.

3.3.2 Comparison between OSER and the synthetic compartments

Initial characterisation of the G22Y compartment by Ian Moore's group suggested that it differed from OSER in some key properties, most notably the compartment morphology,

and that usually only a single structure was formed, unlike for OSER[155,156]. With the more in-depth characterisation of this phenomenon in this chapter, we can now use more properties to compare it with OSER to get a better understanding of the relationship of the two.

Morphologically, G22Y compartments with their helical organisation (Figure 3.6A & Supplementary Figure S4) are unlike known OSER structures. C22Y presents globular compartments that are broadly similar to some previously described OSER structures[148]. The fibrous structure of 22Y sometimes resembles OSER lamellae, but these intertwining membrane fibres often form a more complex, globular structure (Figure 3.6B), unlike lamellae which generally remain in an elongated shape[143]. However, striking similarity can be observed between the OSER structure induced by the *nuc* mutation in the MVP1 protein in Arabidopsis[151] and the 22Y structures in this study. MVP1 is not predicted to be oligomerising or membrane-bound and was shown to be indirectly involved in ER-to-Golgi protein transport[150,151,181,182]. It was suggested that MVP1 mutations induce OSER by disrupting ER export and cause aberrant retention of large quantities of other OSER-inducing membrane proteins in the ER[143]. The morphological similarity between the MVP1 induced OSER and the 22Y compartments raise the interesting possibility that ER-to-Golgi transport might be blocked in cells overexpressing 22Y. However, since AnalyzER results showed that the ER network dynamics remain unchanged, this seems less likely (Figure 3.12). An alternative explanation could be that the aberrantly retained proteins in MVP1 induced OSER are similar to 22Y, thus causing the striking resemblance. It is important to note, that the swollen ER bodies observed by overexpressing 22Y were never before described in connection to OSER.

The formation of OSER did not block normal ER function and plant development[148], which is in line with the findings above; G22Y expressing cells retained the compartments for up to six weeks and remained viable. The AnalyzER analysis also confirmed that the rest of the ER network remains mostly undisturbed, and thus functional impairment is unlikely. FRAP results also confirmed that the compartments retain a connection to the ER, similar to OSER. The retarded lumenal recovery rate (Figure 3.9A) is also consistent with previous OSER FRAP experiments, which suggested that diffusion of proteins into the OSER structures is limited[150].

The proteins inducing OSER are diverse, but fall in two general categories: either indirectly inducing OSER by disrupting the ER-to-Golgi export system (such as MVP1) and probably aberrantly retaining OSER-inducing membrane proteins; or membrane proteins that directly induce OSER[143]. My experiments also showed that a transmembrane domain was essential to induce the formation of the ER-derived compartments. The necessity of a cytosolic oligomerising domain for OSER induction is more controversial, with experimental data supporting this, but also recent findings showing that in some cases the overexpression of an ER-resident transmembrane domain is sufficient to induce OSER on its own[143,148,149]. This is in contrast with my findings, where a cytosolic, antiparallel oligomerising domain was essential for compartment formation.

Overall, there seems to be significant similarities between the described compartment and OSER, but also notable differences in some cases in the overall morphology and in the structure of the formation-inducing constructs as well. As such, while it cannot be excluded that the compartments are a special subtype of OSER, it is also not definitely proven.

3.3.3 Potential applications of the synthetic compartment

This initial exploration of using self-organising, oligomerising polyproteins to generate an ER-derived compartment has shown great promise, and characterisation of the different compartment types suggested that a wide range of potential applications might be possible, similar to what was previously suggested for OSER[31,143].

The modularity of the polyproteins enables the attachment of additional elements to the inner or outer surface of the compartments. One possible application stemming from this would be to use the compartment for sequestration of high concentration of valuable proteins, especially toxic proteins (e.g., OmpG porins) from the cytosol and the rest of the ER, by immobilising the target proteins by the scaffold, thus the bulk of the cell is shielded from their effects, especially if they are captured on the ER lumenal side without entering the cytosol. Capture in the ER lumen would also be beneficial for commercially valuable targets which face degradation by proteases further downstream, such as antibodies, which could be immobilised in the luminal side by a genetically fused Protein A[183]. Recovery of proteins via purification of the protein scaffold would require genetically fused affinity tags (e.g., a his-tag), but alternatives could be investigated as well, such as microsomal preparations or sucrose gradient-based solutions. A successful proof of principle study of this has been carried out using mRFP as a genetically fused lumenal cargo by Ian Moore's group[156]. Beyond direct genetic fusion of valuable cargo proteins, the introduction of a more versatile 'plug-and-play' type of binding domain, such as SpyCatcher would enable a single universal scaffold to be used with a variety of targets (that have the appropriate SpyTag domain which forms a covalent bond with SpyCatcher)[184,185], significantly reducing

the time needed to establish and optimise the expression of new products. More work on this topic is described in detail in Chapter 5.

Another possible functionalisation of the compartment is to utilise the structure to create a synthetic microdomain. This is possible, because FRAP experiments have shown that the compartment forming proteins in the membrane are immobilised in close proximity to each other (although it is unclear if the other membrane proteins present are immobilised: this could be uncovered by e.g., FRAP of TAR2-RFP). Furthermore, exchange between trapped cytosol and bulk cytosol, as well as the lumen of the compartment and the lumen of the ER is slowed (Figure 3.9). These properties together fulfil the main criteria for a probabilistic microdomain, where high local concentrations of sequential enzymes in a pathway, combined with the slower rate of flow into the bulk aqueous solution, increase the chance of an intermediate substrate to encounter the subsequent enzyme in the pathway (compared to leaving the microdomain and mix into the bulk aqueous solution)[83]. The relevant enzymes could be attached by either genetic fusion or using the modular SpyCatcher-SpyTag system[184,185], and depending on the requirements of the pathway (i.e., pH, salt concentration, co-factors, etc.) the microdomain can be assembled on either the cytosolic or lumenal side of the compartment. Preliminary work on this topic is described in detail in Chapter 5.

Chapter 4: Further investigation into the relationship between the structure of the synthetic compartments and the constitutive domains of the compartment-forming protein constructs

4.1 Introduction

In Chapter 3, I established that using a variety of oligomerising membrane proteins derived from previous work by Moore *et al.* I can generate many different types of self-organising ER-derived compartments. Based on the range of compartment morphologies and the structure of the corresponding protein constructs I have also proposed a hypothesis as to how the orientation of the oligomerising domains in these proteins may affect the structure of the resulting compartment. Finally, during the course of my in-depth characterisation of the compartments, I have been able to put forward a number of potential applications for the synthetic compartment, including its use as a way to improve production of difficult-to-express membrane proteins and to serve as the basis for a synthetic microdomain for probabilistic metabolite channelling.

However, testing these claims and hypotheses required the development and characterisation of a substantial number of additional constructs. In order to test the hypothesis regarding the relationship between compartment morphology and the structure of the oligomerisation domains of the protein constructs, several new polyproteins with a range of component orientations were assessed to determine if these followed the proposed rules. With a greater range of compartment-forming protein constructs, and a deeper understanding of the formation and structure of the compartments, I hoped that I would also be able to gain further insight into the question

of whether these compartments are truly distinct from OSER or form a special, previously undescribed, subcategory of OSER.

These new 'second-generation constructs' were designed with the addition of extra peptide tags that would serve as the basis for functionalisation of the compartments. One of the proposed applications – construction of a metabolic microdomain - required the ability to covalently bind other proteins or enzymes to the scaffold proteins *in vivo*. For this, the SpyTag peptide[184,185] was selected as an appropriate cytosol-facing tag that would allow the covalent attachment of any cytosolic proteins to the surface of the compartment that have the corresponding SpyCatcher domain[184,185]. To use the compartment for protein production required the addition of an affinity-purification tag to allow the rapid, specific, and high-yield recovery of expressed proteins of interest. The well-established 6xHis tag was selected for this purpose[186].

To ascertain that the addition of these new tags did not alter the compartments' other properties that made them attractive for these potential applications in the first place, the new constructs not only had to be expressed, but also characterised in a similar fashion to that described in Chapter 3. If compartments with the required tags retained their properties (such as compartment size, diffusion barrier to the cytosol, construct modularity, and little effect on the bulk ER), work would then proceed to testing the proposed functionalisations.

4.2 Results

4.2.1 Design of the second-generation compartment forming protein constructs

In order to expand on the results from Chapter 3 and gain further insight into the relationship between the structure of the protein constructs and the formed compartments, a number of new, second-generation protein constructs were designed and developed (Figure 4.1). For the purposes of potential functionalisations discussed in Chapter 3.3.3, two C-terminal tags were attached to most second-generation constructs: a 6xHisTag for efficient purification of the proteins, and a SpyTag to allow covalent binding of other proteins with the appropriate SpyCatcher domain[184,185].

These additions onto the C22Y proteins yielded the C22Y-SpyT-his construct, and the two derivatives C22Y-ΔSpyT-his and C22Y-SpyT-Δhis, with either of the tags removed. The effect of the length and rigidity of linkers between the different protein domains of the construct was also investigated, by introducing a long stiff linker $((AP)_5)$[174] between the YFP and the SpyTag domains in C22Y-SpyT-Δhis to build C22Y-SL-SpyT, and by removing the linker altogether in C22Y-NL-SpyT, and in a new 22Y derived construct by introducing a long flexible linker $((SGGGG)_3ELDEESILKNE)$[174,175] between the BP22 transmembrane domain and the cytosolic YFP, to form 22LLY-SpyT-his (Figure 4.1).

Since the orientation of the dimerizing domains was hypothesised to have great effect on the morphology of the compartment in Chapter 3, a number of constructs were built with differently orientated dimerizing proteins to test this. Two new constructs were designed in which the previously used synthetic parallel dimerizing coiled-coil domain CC-Di was

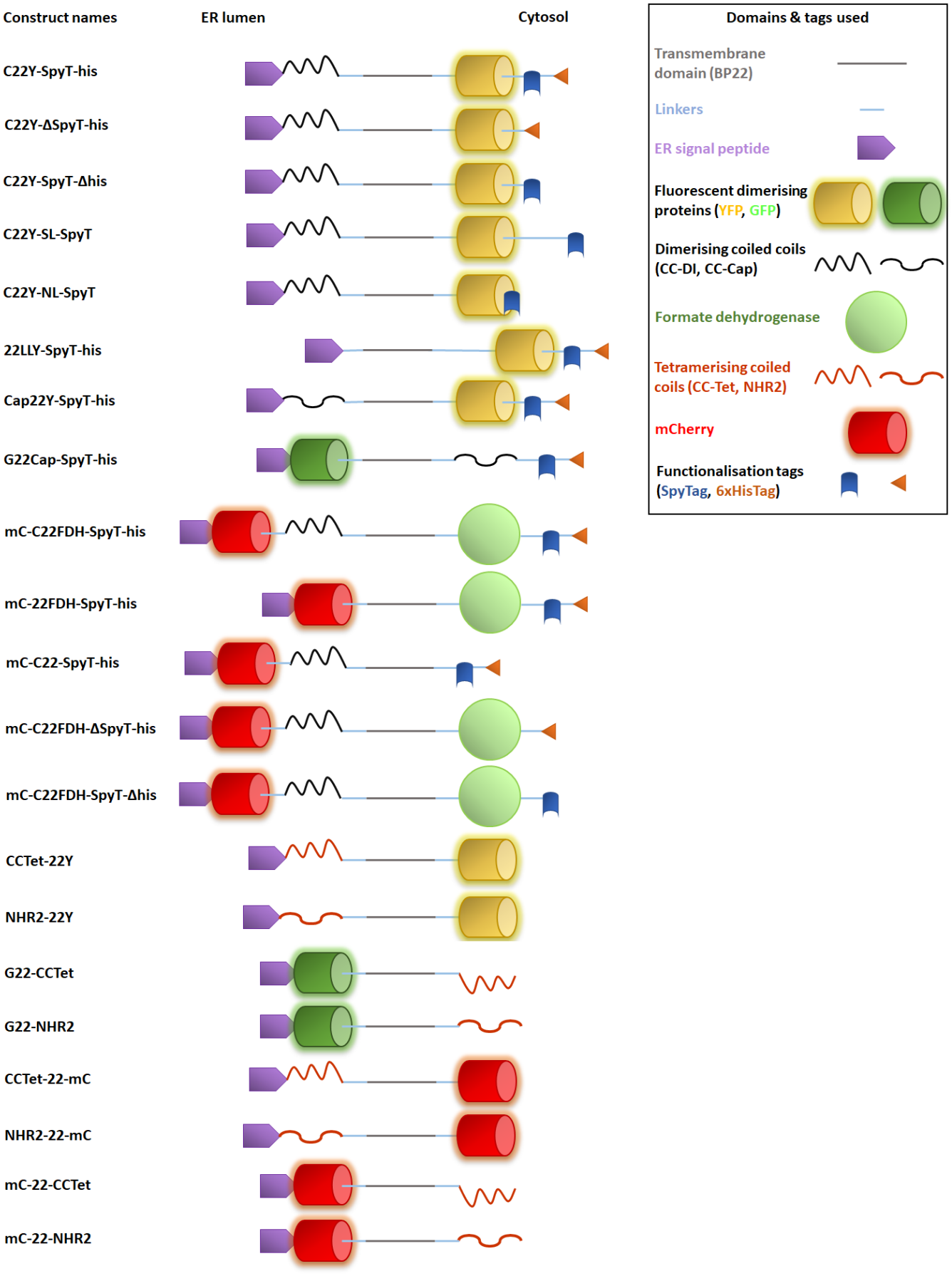
replaced by the antiparallel dimerizing coiled-coil domain CC-ap from bovine IF1[172], yielding Cap22Y-SpyT-his and G22Cap-SpyT-his (Figure 4.1).

Functionalisation of the compartments for metabolic engineering would require the attachment of an enzyme to the surface of the compartment. A modular approach to this is the aforementioned SpyCatcher – SpyTag system, but an alternative is the direct genetic fusion of an enzyme to the cytosolic C-terminus of the construct. To reduce the overall size of the protein, I wanted to test if this fused enzyme could serve as the oligomerization domain as well, by replacing the cytosolic YFP in C22Y-SpyT-his with the antiparallel dimerising enzyme formate dehydrogenase (FDH) from *S. tuberosum*[187]. Since this meant the removal of the YFP, the monomeric red fluorescent protein mCherry was attached to the N-terminus of the construct to enable fluorescent imaging of the resulting compartments. This construct (mC-C22FDH-SpyT-his) had a number of derived variants, to test the effects of the removal of either dimerizing domain (mC-22FDH-SpyT-his and mC-C22-SpyT-his), and the C-terminal tags (mC-C22FDH-ΔSpyT-his and mC-C22FDH-SpyT-Δhis) (Figure 4.1).

All previous constructs utilised dimerizing domains to form compartments, but I wanted to also test higher order interactions, and their effects on the ultrastructure of compartments, to determine if these constructs would follow the established hypothesis on the relationship between compartment and construct structure (Chapter 3.3.1). To this end, a number of tetrameric constructs were designed, using two tetrameric domains: the synthetic parallel homotetrameric CC-Tet[171], and the antiparallel homotetrameric NHR2 from *M. musculus*[173]. These were introduced on both the lumenal (NHR2-22Y and CCTet-22Y) and cytosolic side (G22-NHR2 and G22-CCTet). A variant was also developed for each

where the fluorescent dimerising YFP or GFP were replaced by mCherry (NHR2-22-mC, CCTet-22-mC, mC-22-NHR, mC-22-CCTet) (Figure 4.1) to mirror the experiments conducted in Chapter 3 allowing direct comparison to the original constructs and results presented there.

Figure 4.1 (below): Schematic of the second-generation compartment forming polyprotein constructs investigated. All constructs contain the transmembrane domain BP22 (grey), linkers (light blue) and an ER signal peptide (purple) which targets the polyproteins to the ER membrane. Oligomerising domains used were GFP (dark green), YFP (yellow), dimerizing coiled coil domains (black) CC-Di (synthetic, parallel dimerising) and CC-ap (from bovine IF1, antiparallel dimerising), formate dehydrogenase (FDH, from *S. tuberosum,* antiparallel dimerising, light green) and tetrameric coiled coil domains (brown) CC-Tet (synthetic, parallel dimerising) and NHR2 (from *M. musculus,* antiparallelly dimerising). Constructs with no fluorescent dimerizing domains had a monomeric mCherry (red) attached to them to allow visualisation. To test the effects of C-terminal tags required for functionalisation of the compartments, SpyTag (dark blue) and 6xHisTag (orange) were also introduced to a number of constructs.


Construct names
ER lumen
Cytosol
C22Y-SpyT-his
C22Y-ΔSpyT-his
C22Y-SpyT-Δhis
C22Y-SL-SpyT
C22Y-NL-SpyT
22LLY-SpyT-his
Cap22Y-SpyT-his
G22Cap-SpyT-his
mC-C22FDH-SpyT-his
mC-22FDH-SpyT-his
mC-C22-SpyT-his
mC-C22FDH-ΔSpyT-his
mC-C22FDH-SpyT-Δhis
CCTet-22Y
NHR2-22Y
G22-CCTet
G22-NHR2
CCTet-22-mC
NHR2-22-mC
mC-22-CCTet
mC-22-NHR2
Domains & tags used
Transmembrane domain (BP22)
Linkers
ER signal peptide
Fluorescent dimerising proteins (YFP, GFP)
Dimerising coiled coils (CC-DI, CC-Cap)
Formate dehydrogenase
Tetramerising coiled coils (CC-Tet, NHR2)
mCherry
Functionalisation tags (SpyTag, 6xHisTag)

4.2.2 Initial characterisation of the second-generation constructs

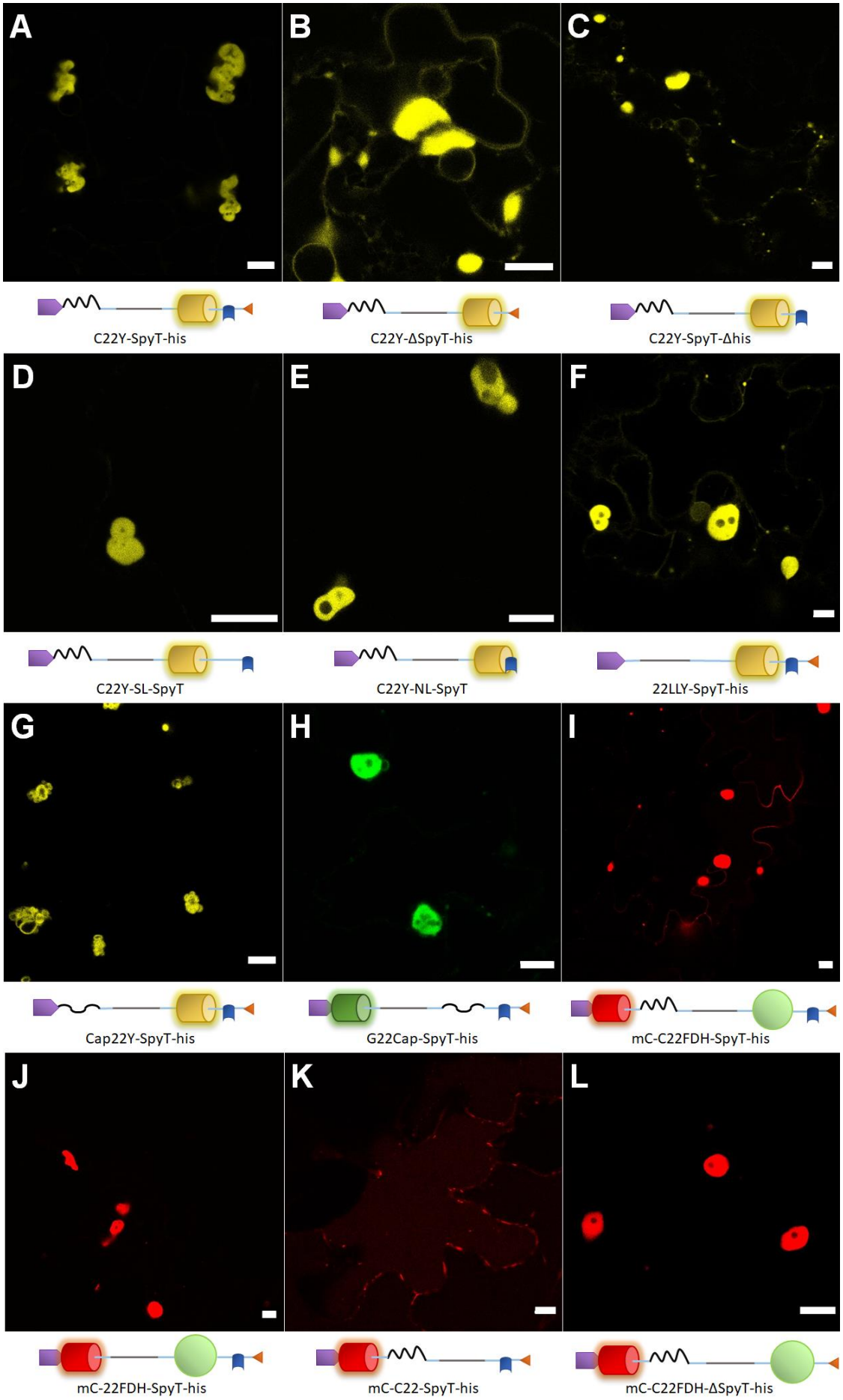
All second-generation constructs (Figure 4.1) were successfully cloned (confirmed by Sanger sequencing of plasmids isolated from each construct-carrying strain) and transiently expressed in mature tobacco leaves using *A. tumefaciens*-mediated transformation. Protein expression was confirmed by confocal fluorescence microscopy (Figure 4.2) and a range of compartment morphologies were observed.

Most constructs (C22Y-SpyT-his, C22Y-ΔSpyT-his, C22Y-SpyT-Δhis, C22Y-SL-SpyT, C22Y-NL-SpyT, 22LLY-SpyT-his, G22Cap-SpyT-his, mC-C22FDH-SpyT-his, mC-22FDH-SpyT-his, mC-C22FDH-ΔSpyT-his, mC-C22FDH-SpyT-Δhis, CCTet-22Y, NHR2-22Y) formed a single large (10-35 µm diameter) compartment in transformed cells (Figure 4.2A-F, H-J, L-O). These closely resembled the large G22Y and C22Y compartments previously described in Chapter 3, although notably many (but not all) of these compartments were punctuated by large numbers of small (1-2 µm) and/or a few moderately large (3-5 µm) spherical holes without any fluorescent signal, which was not typical of the aforementioned G22Y and C22Y compartments. A notable exception to this morphology was produced by Cap22Y-SpyT-his, where the larger compartments (10-20 µm) were composed of easily distinguishable randomly curved strands (Figure 4.2G). Compartments remained stable for several weeks after transient transformation, similar to what was observed for first generation constructs (Figure S2).

Some constructs failed to produce large compartments, instead showing strong fluorescent signal in the ER and in smaller spherical puncta (mC-C22-SpyT-his, G22-CCTet, G22-NHR2, CCTet-22-mC, NHR2-22-mC, mC-22-CCTet, mC-22-NHR2) (Figure 4.2K, P-U) resembling images of G22 and G22C from Chapter 3. These constructs, when containing an

mCherry domain as well, often showed a vacuolar red fluorescent signal (see Figure 4.2K, T-U for examples) which grew in strength over time, especially beyond 7 days after transformation, suggesting an accumulation of the construct proteins in the vacuole.

To determine if the aforementioned novel compartment structure (large compartment resembling the C22Y compartment but with several spherical holes) is derived from the ER similarly to all other previously described compartments, C22Y-SpyT-his was co-expressed with ER lumenal marker RFP-HDEL and ER membrane marker TAR2-RFP (Figure 4.3). The two ER makers showed clear overlap with the fluorescent signal from the compartments, confirming that the membrane and lumen of the new compartment type is derived from the ER. RFP-HDEL was also co-expressed with the tetrameric constructs CCTet-22Y and NHR2-22Y to confirm their similarity to previously described compartments, with ER luminal and compartment signals overlapping as well (Figure 4.3). When the tetrameric constructs G22-CCTet and G22-NHR2 were co-expressed with RFP-HDEL, the two signals clearly separated, with no RFP fluorescence detected in the small spherical compartments (Figure 4.3), suggesting that these construct form a morphology similar to that of G22 and G22C in Chapter 3, and were not further investigated.


A
B
C
C22Y-SpyT-his
C22Y-ΔSpyT-his
C22Y-SpyT-Δhis
D
E
F
C22Y-SL-SpyT
C22Y-NL-SpyT
22LLY-SpyT-his
G
H
I
Cap22Y-SpyT-his
G22Cap-SpyT-his
mC-C22FDH-SpyT-his
J
K
L
mC-22FDH-SpyT-his
mC-C22-SpyT-his
mC-C22FDH-ΔSpyT-his

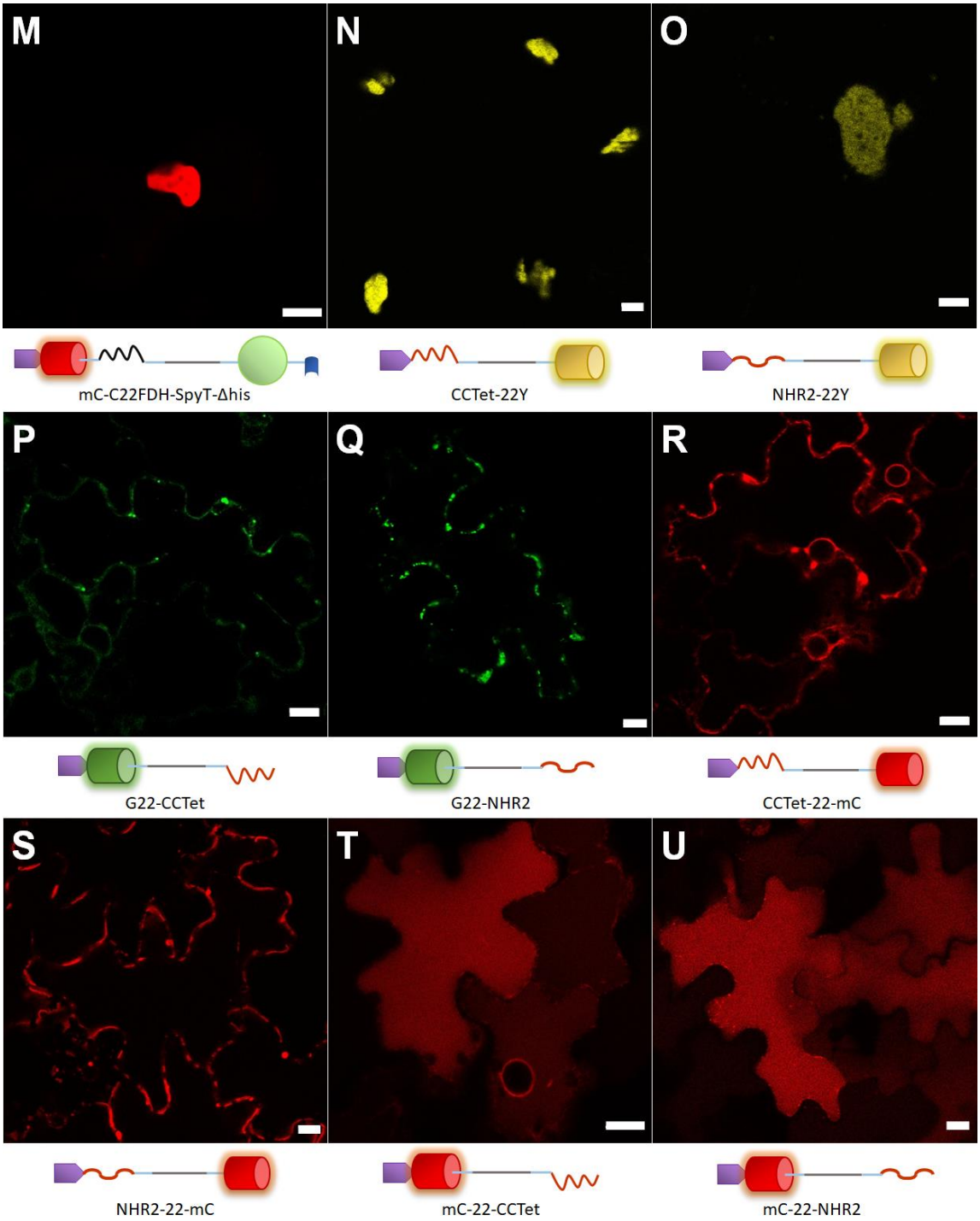
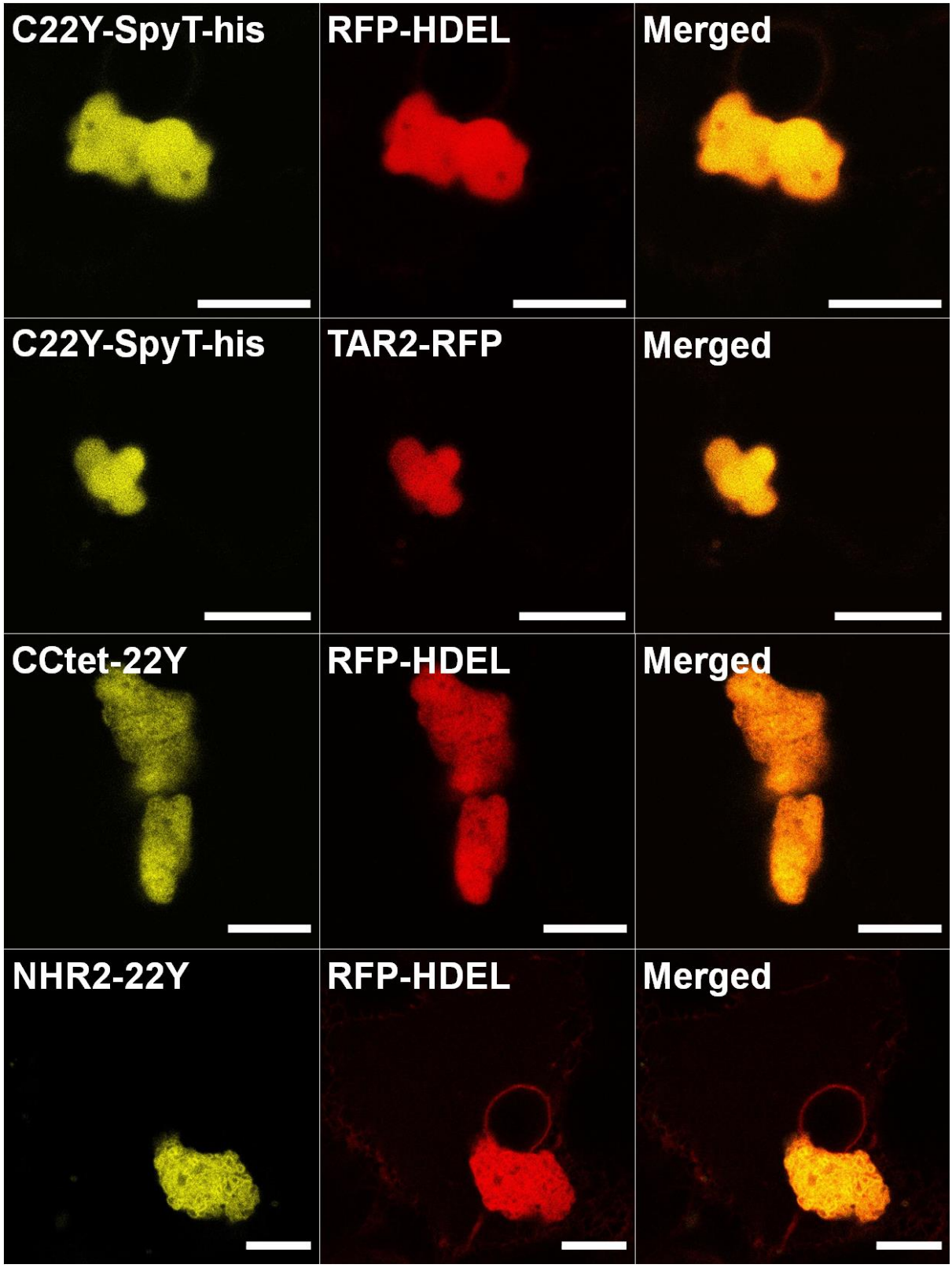


Figure 4.2 (above and previous page): Confocal microscopy of the second-generation constructs 5-7 days after transformation of mature tobacco leaves. Green: GFP. Yellow: YFP. Red: mCherry. The schematic of the corresponding construct as in Figure 4.1 is depicted under each image. A-F, H-J, L-O show large compartments usually with small spherical puncta devoid of fluorescent signals. G shows a notably different compartment morphology, with aggregated randomly curved strands. K and P-U formed no large compartments, but show dispersed signal in the ER, in small spherical compartments and often in the vacuole. Scale bars: 10 µm.


C22Y-SpyT-his
RFP-HDEL
Merged
C22Y-SpyT-his
TAR2-RFP
Merged
CCtet-22Y
RFP-HDEL
Merged
NHR2-22Y
RFP-HDEL
Merged

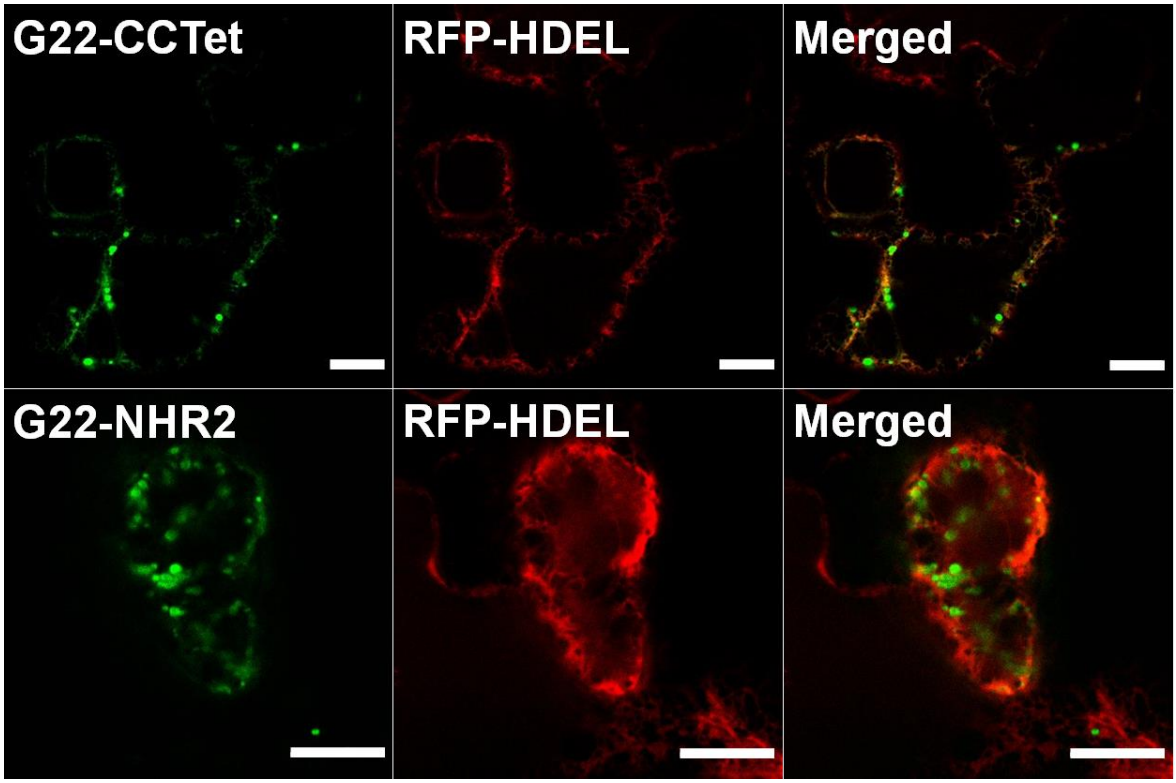
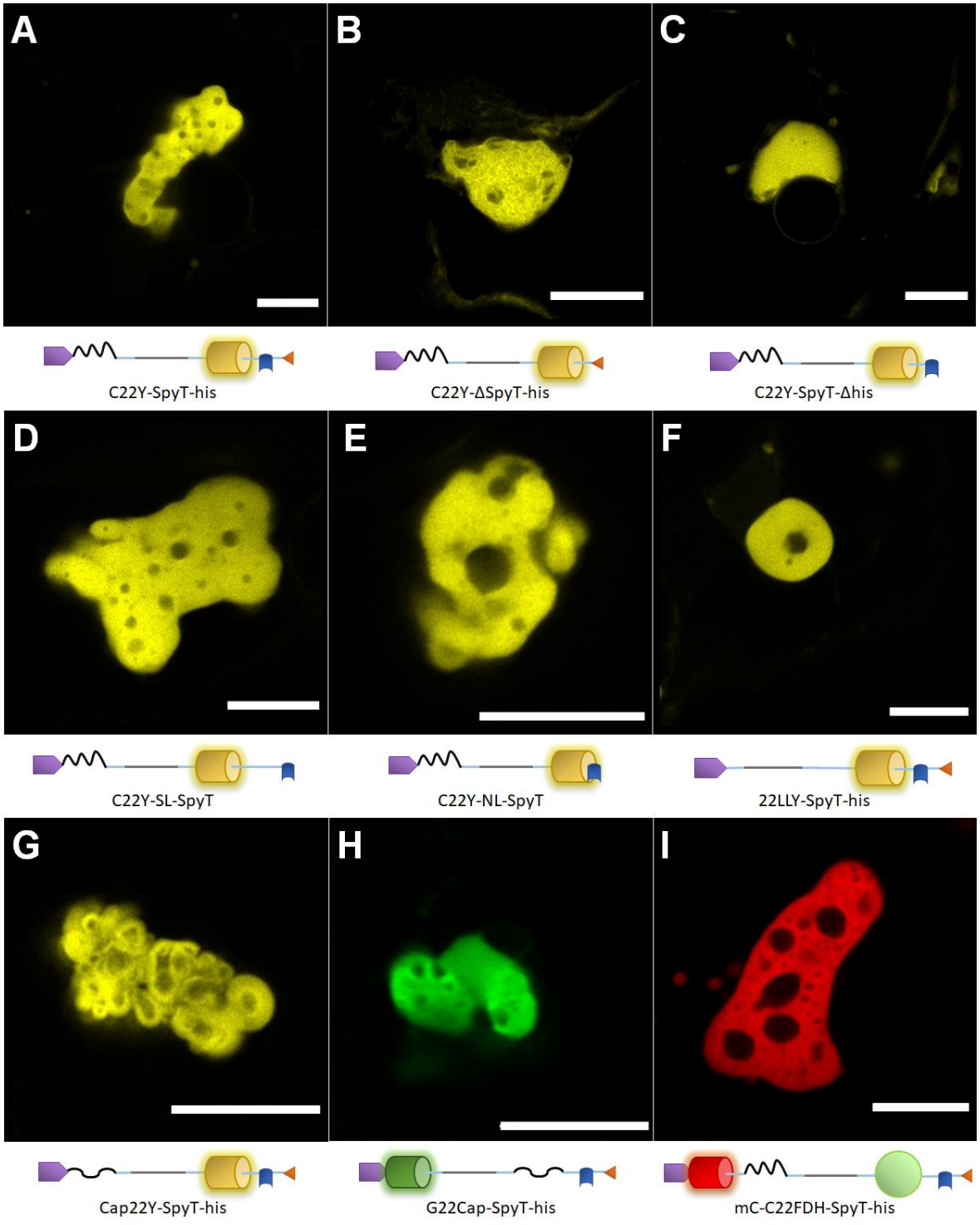


Figure 4.3 (above and previous page): Co-expression of ER markers with the compartments to determine if the lumen and membrane of the compartments are derived from the ER. RFP-HDEL is an ER luminal marker, TAR2-RFP is an ER membrane marker. Signal from large compartment forming constructs (C22Y-SpyT-his, CCTet-22Y & NHR2-22Y - yellow) show overlap with the ER marker signals (red), unlike the constructs producing small spherical compartments (G22-CCTet & G22-NHR2 - green). Images were visualised using fluorescent confocal microscopy of mature tobacco leaves, 7 days after transformation. Scale bars: 10 µm.

4.2.3 High-resolution analysis of second-generation compartments

The constructs forming large compartments were further investigated using higher resolution airy-scanning confocal microscopy, to allow comparison to previously described compartments (Figure 4.4). Closer inspection of these constructs confirmed the presence of a new compartment morphology, which was similar to the C22Y compartments (even coverage of fluorescent signal which could not be further resolved by fluorescent airy-scanning confocal microscopy) (Figure 3.6C) but with numerous spherical puncta (Figure 4.4A-F, H-L). Notably, the second-generation constructs which formed large compartments and had at least one C-terminal tag attached almost exclusively produced this new structure (C22Y-SpyT-his, C22Y-ΔSpyT-his, C22Y-SpyT-Δhis, C22Y-SL-SpyT, C22Y-NL-SpyT, 22LLY-SpyT-his, G22Cap-SpyT-his, mC-C22FDH-SpyT-his, mC-22FDH-SpyT-his, mC-C22FDH-ΔSpyT-his, mC-C22FDH-SpyT-Δhis). It is likely, that these structures are below the resolution limit of airy-scanning confocal microscopy, similar to C22Y compartments. However, to conclusively determine this, further TEM imaging is recommended in the future, which was not possible to carry out during the course of this project, due to constraints imposed by the Covid-19 pandemic.

The only exception to this rule was discovered by higher resolution images of Cap22Y-SpyT-his (Figure 4.4G), which revealed curved strands closely resembling G22Y (Figure 3.6A). Notably NHR2-22Y (Figure 4.4N), which does not have a C-terminal tag, showed a similar morphology as well, albeit with less ordered edges. CCTet-22Y (Figure 4.4M) resembled the structure of C22Y most closely (Figure 3.6C).


A
B
C
C22Y-SpyT-his
C22Y-ΔSpyT-his
C22Y-SpyT-Δhis
D
E
F
C22Y-SL-SpyT
C22Y-NL-SpyT
22LLY-SpyT-his
G
H
I
Cap22Y-SpyT-his
G22Cap-SpyT-his
mC-C22FDH-SpyT-his

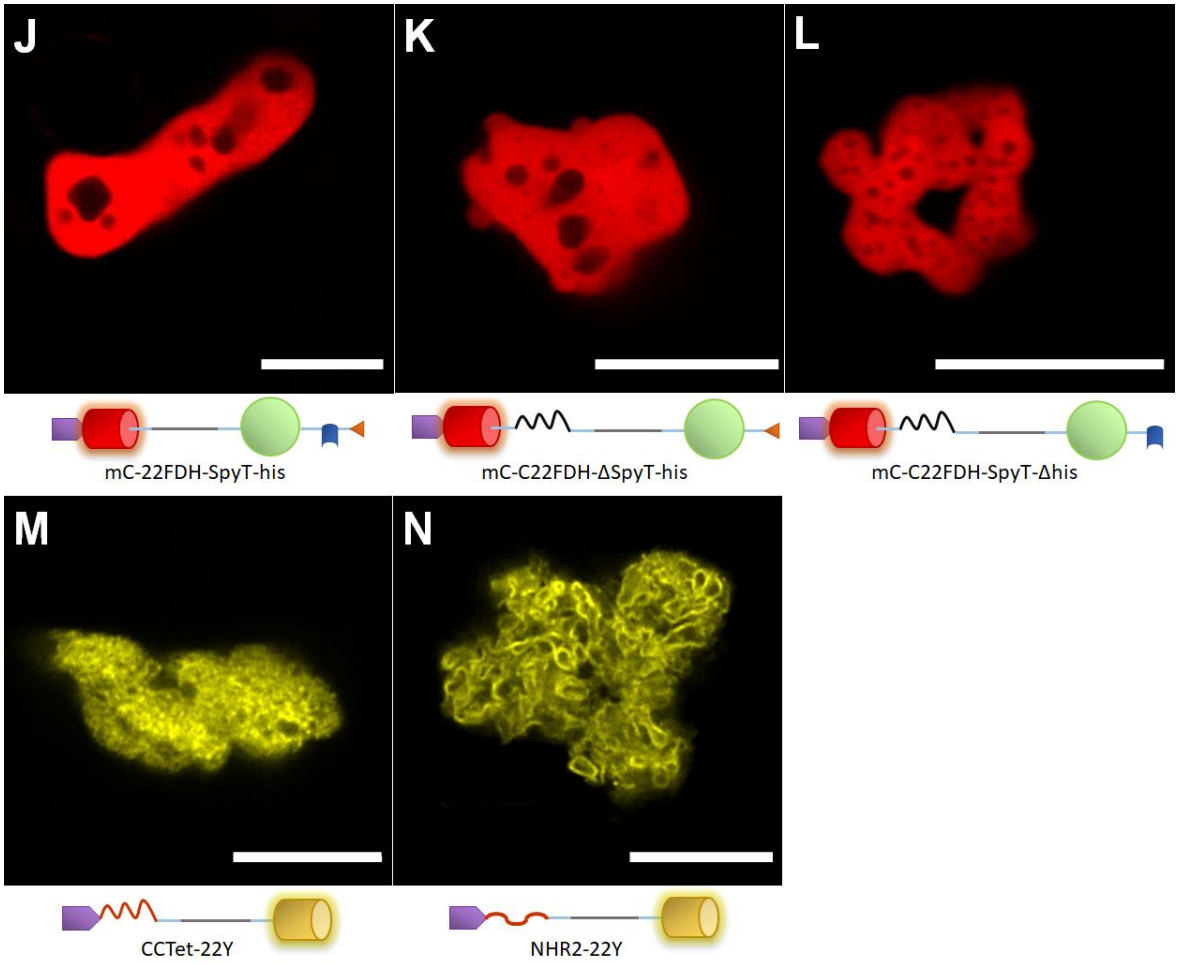


Figure 4.4 (above and previous page): High-resolution airy-scanning confocal microscopy images of second-generation compartment forming constructs 5-7 days after transformation of mature tobacco leaves. Yellow: YFP. Green: GFP. Red: mCherry. The schematic of the corresponding construct from Figure 4.1 is depicted under each image. A-F, H-L: A distinct, new structure can be observed which is similar to previously described C22Y compartments (Figure 3.6C), but with numerous small-, and/or a few moderately large holes in the structures. M: These holes are not apparent, with this construct showing an identical morphology to C22Y. G&N: Individual strands are more pronounced in these compartments with curved edges. While both resembles G22Y compartments (Figure 3.6A), N shows a more random organisation of the strands. Scale bars: 10 µm.

4.2.4 Compartment diffusion dynamics

One key objective for the functionalization of the compartments was to develop them as a synthetic metabolic microdomain to promote probabilistic metabolite channelling. The SpyTag – SpyCatcher system was added for this purpose to allow easy binding of the required enzymes to the cytosolic surface of the compartments. Since the addition of a C-terminal SpyTag (and/or 6xHisTag) almost universally lead to the formation of a new compartment type, it was essential to investigate if it was still a viable candidate as a platform for a synthetic microdomain.

In order to assess this, mature tobacco leaves transformed by mC-C22FDH-SpyT-his were infiltrated with FDA, a cytosolic yellow fluorescent small molecule probe. Compartments were then photobleached (FRAP) and recovery of FDA into the cytosol trapped between the membranes of the compartments was measured. The recovery rate of FDA was significantly retarded compared to the rest of the cytosol (Figure 4.5). This suggested that the cytosol trapped between the membranes of the compartment is still connected to the bulk cytosol, but there is a notable diffusion barrier even for small molecules, confirming that the new compartment is still a viable candidate for the development of a synthetic microdomain.

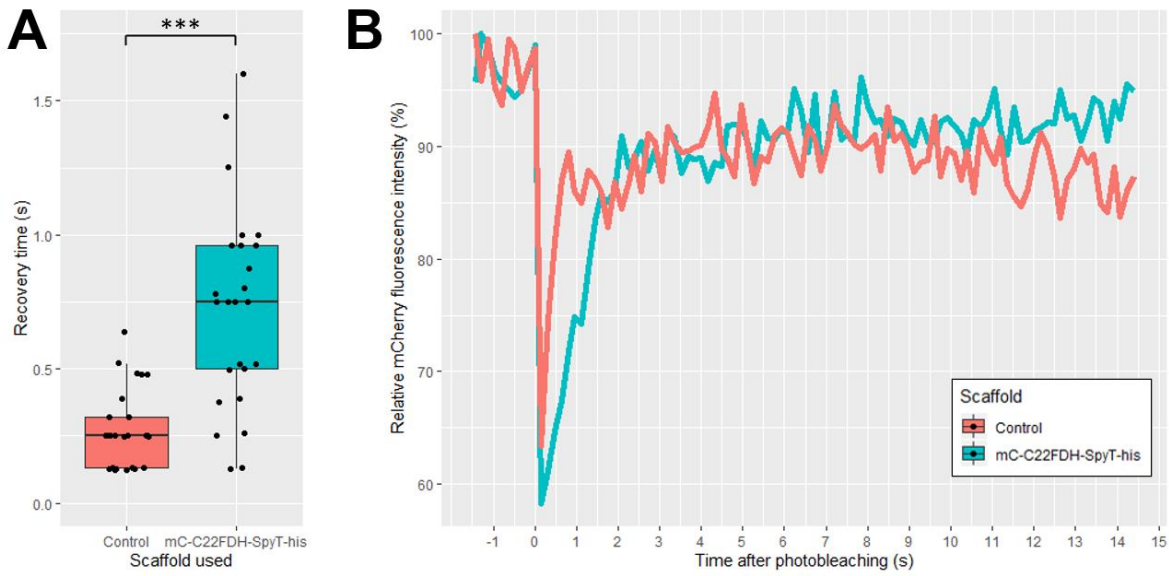


Figure 4.5: Retarded cytosolic fluorescence recovery rate of the small molecule probe FDA in compartments after FRAP. Red: WT control (FDA only); Blue: mC-C22FDH-SpyT-his A: Recovery time is significantly longer (t-test, $p < 10^{-7}$) for mC-C22FDH-SpyT-his compared to the control. n = 25 for both categories. The boxes represent the interquartile range, the horizontal line in the box shows the median, and the whiskers the minimum and maximum values (excluding outliers). B: Representative example of fluorescence signal recovery after photobleaching. Bleaching occurs at t = 0 sec. Fluorescence intensity is normalised to the maximum value of each collected image series.

4.2.5 ER network dynamics

To determine if the presence of the compartments had a significant effect on the host cells, and to assess if the rest of the ER was disturbed, the ER network dynamics were captured using short (20 seconds) time-series images that were quantitatively analysed using the AnalyzER software package[168]. To keep the analysis consistent and comparable with previous results from AnalyzER analysis of G22Y, 22Y and C22Y (Chapter 3.2.5), the same 41 ER parameters were investigated using the same computational pipeline as before (Table 3.1). These include topological, morphological and kinetic properties of the tubules, cisternae and polygonal regions enclosed by the ER.

Two separate AnalyzER analyses were carried out on selected second-generation constructs. The first focused on the new compartment morphology observed when a C-terminal tag was attached to the compartment forming constructs. This analysis utilised C22Y-SpyT-his, co-expressed with RFP-HDEL (the latter was visualised to acquire images of the ER network). The AnalyzER data was subjected to PCA, and the principal components were individually assessed for any statistically significant differences (using MANOVA and ANOVAs) between compartment expressing cells and controls (Figure 4.6). None of the principal components showed a statistically significant difference ($p>0.05$), suggesting that the compartments' effects on the rest of the ER network was relatively insignificant, and this new compartment is also not expected to have a major impact on the expressing cells.

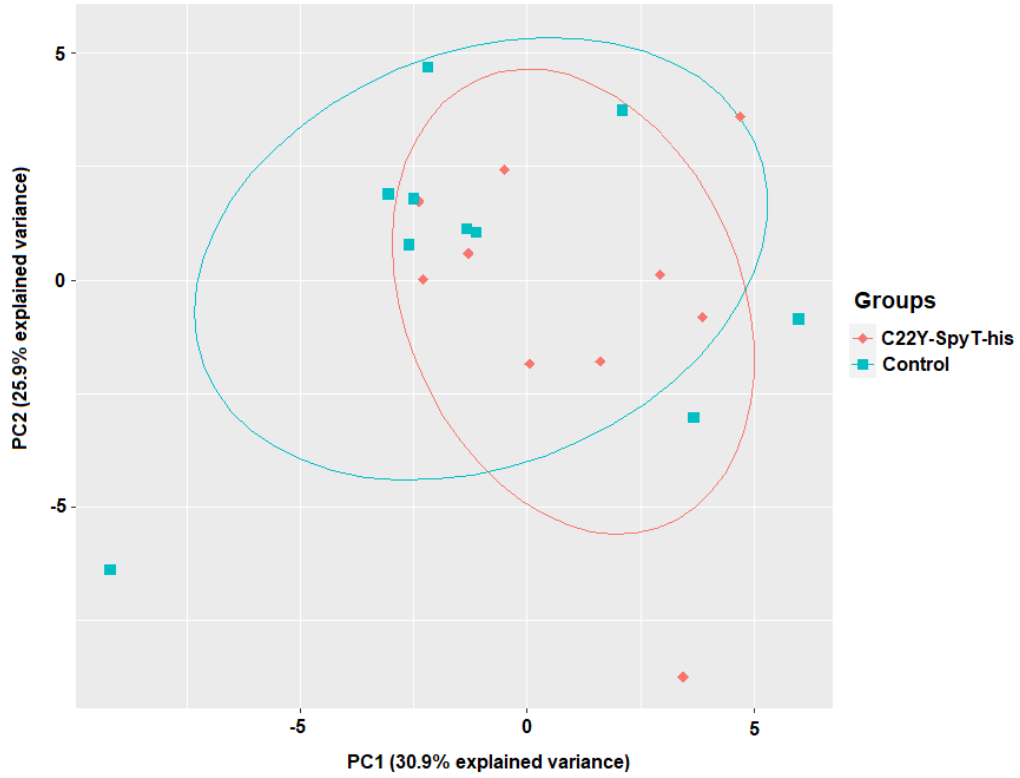


Figure 4.6: Principal component analysis of the effects of C22Y-SpyT-his on ER network dynamics. The two primary principal components are shown here, cumulatively accounting for 56.8% of the variance in the data. C22Y-SpyT-his and WT (control) values overlap significantly, confirming the statistical analysis which found no significant differences between the two groups. n = 10 for both groups.

The second analysis focused on the effects of the compartment forming tetrameric constructs NHR2-22Y and CCTet-22Y on the ER network, testing if the higher-order coiled coil interactions could have effects or changes that are not immediately apparent using visual inspection of high-resolution confocal microscopy (Figure 4.4M-N). The AnalyzER data was statistically analysed using the same *in silico* pipeline, and this showed statistically significant differences between the groups for the first principal component (PC1, explaining 31.5% of the total variance, $p = 0.0029$), and the Tukey's HSD test confirmed a significant difference ($p = 0.002$) for PC1 between NHR2-22Y and the controls (Figure 4.7).

To determine the exact effect of the tetrameric constructs on the ER, the 12 parameters which contributed the greatest effect size to PC1 were further investigated using individual ANOVAs, with 3 parameters retaining significance (Table 4.1): cisternal mean distance ($p = 0.039$), cisternal max distance ($p = 0.028$) and cisternal area ($p = 0.033$). This suggests that the tetrameric construct NHR2-22Y indeed disrupted the rest of the ER to a significant enough degree to be detected. Notably, all three significant parameters describe morphological properties of the ER cisternae, especially the size, shape and area of the cisternae. However, none of the measured kinetic parameters, or parameters of the tubules or the polygons enclosed by the ER were significantly altered.

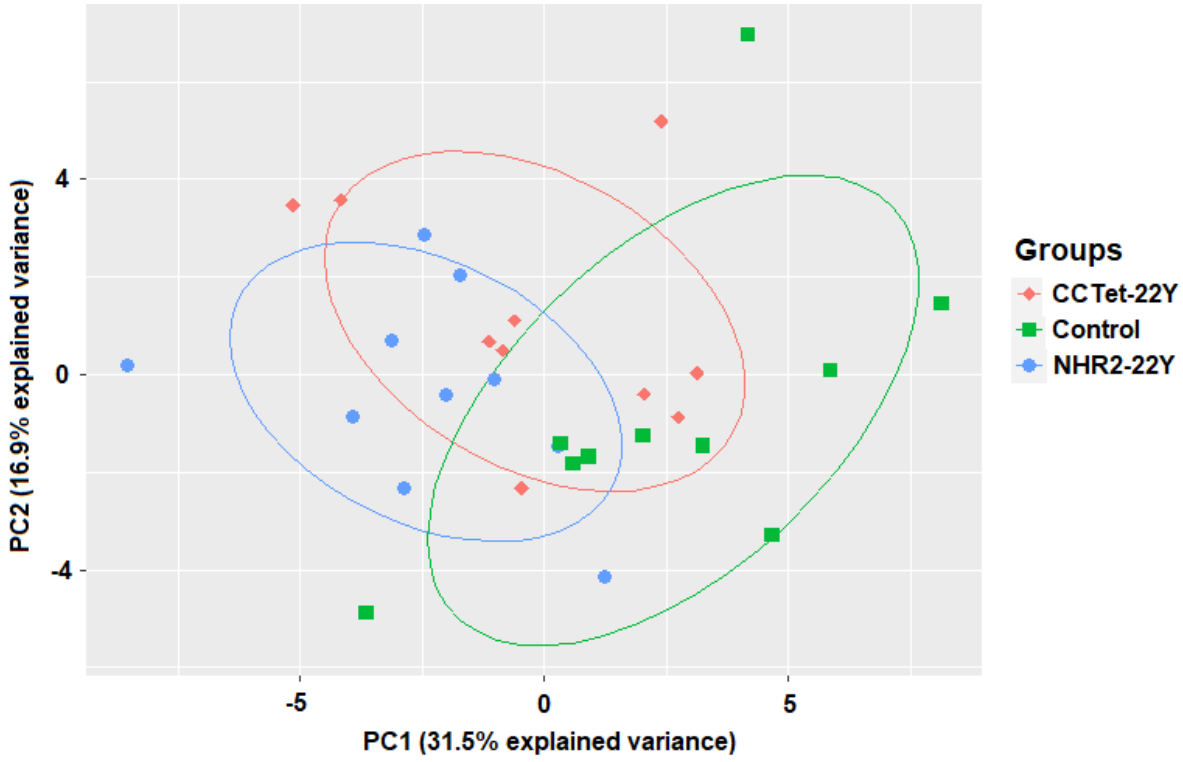


Figure 4.7: Principal component analysis of the effects of the tetrameric constructs on ER network dynamics. The two primary principal components are shown here, cumulatively accounting for 48.4% of the variance in the data. The tetrameric constructs and the control values overlap considerably, but statistical analysis shows that control and NHR2-22Y is significantly different for PC1 ($p = 0.002$). $n = 10$ for all groups.

Parameter type	Parameter name	p-value	Significance
Cisternae	cisternal_Circularity	0.999	
Cisternae	cisternal_MaxPersistency	0.710	
Cisternae	cisternal_MeanDistance	0.039	*
Cisternae	cisternal_MaxDistance	0.028	*
Cisternae	cisternal_Node_Degree	0.222	
Cisternae	cisternal_Area	0.033	*
Polygons	polygon_Area	0.999	
Polygons	polygon_MeanDistance	0.999	
Polygons	polygon_MaxDistance	0.999	
Tubule edges	Flow_coherence	0.678	
Tubule edges	Resistance	0.084	
Tubule edges	Length	0.243	
Tubule edges	Width_center	0.963	
Tubule nodes	node_Omin_Omid	0.193	
Tubule nodes	node_Omid_Omaj	0.062	
Tubule nodes	node_Omin_Omaj	0.054	
Tubule nodes	node_Persistency	0.999	

Table 4.1: The parameters contributing the greatest proportion of variance to PC1 during the AnalyzER analysis of tetrameric constructs. The p-values of individual ANOVAs show that only three parameters show significant difference between the groups (marked with *).

4.3 Discussion

4.3.1 The viability of the second-generation synthetic compartments for potential applications

Two potential applications of the ER-derived compartments were described in Chapter 3.3.3: difficult-to-express or toxic membrane proteins could be sequestered into the compartment to improve protein production, and the compartment could be used as a platform for assembly of a synthetic metabolic microdomain by attaching enzymatic pathways to the cytosolic surface for probabilistic metabolite channelling. Both of these required the addition of further domains to the protein constructs, and the effects of this addition were investigated to ascertain the continued viability of the compartments for these applications.

While the addition of the C-terminal SpyTag and/or 6xHisTag led to the formation of a new type of compartment, the general properties of the compartment did not change significantly: the compartments still consisted of a large, self-organising ER-derived structure, with little effect on the rest of the ER network (Figure 4.6) and having a measurable diffusion barrier between the cytosol trapped between its complex membranes and the bulk cytosol (Figure 4.5). As such, the compartments remain viable candidates for the aforementioned applications.

Furthermore, the work in this chapter demonstrates that other N- or C-terminal domains can be readily replaced or attached to the compartment-forming polypeptide. For example, the addition of the mCherry protein at the N-terminus (for example in construct mC-C22FDH-SpyT-his) allowed the replacement of both of the fluorescent dimerization domains as fluorescent imaging was still possible via the mCherry's red signal. In this

construct, the cytosolic dimerising domain used was the enzyme FDH, which is significantly larger than the fluorescent proteins or coiled coils used previously, but still functioned in the same fashion, and showed that genetically fused enzymes can work as oligomerising domains. This raises the interesting possibility that an oligomerising domain can double as a functional component in the construct as well: for example, if an enzyme from a pathway which is the target of improvement via induction of probabilistic metabolite channelling is a dimer with the monomer pair interacting with each other in an antiparallel orientation, it could potentially be used as both an oligomerising domain and functional enzyme. This would reduce the complexity of the protein constructs, albeit at the cost of modularity. The fact that the tetrameric second-generation constructs generally formed the same compartment structures as their corresponding dimeric variants means that higher-order homo-oligomerising enzymes would be suitable candidates as well, as long as their oligomerising orientation is antiparallel.

4.3.2 The hypothesis of an ER - peroxisome - vacuole pathway for constructs not forming a compartment

Not all constructs formed a large ER-derived compartment when expressed, as shown by G22C and G22 in Chapter 3.2.2, and several second-generation constructs as well, specifically mC-C22-SpyT-his, and six tetrameric constructs (G22-NHR2, G22-CCTet, mC-22-NHR2, mC-22-CCTet, NHR2-22-mC, CCTet-22-mC). Fluorescent imaging of these constructs revealed fluorescent protein in both the unmodified ER (Figure 4.3) and a large number of small spherical compartments, which were identified as peroxisomes for G22C (Figure 3.3). However, further co-expressions of these constructs (other than G22C) with peroxisomal markers would be recommended to conclusively identify their subcellular localisation.

One possible explanation is that elements of these constructs resembled peroxisomal targeting peptides. However, sequence analysis of the tetrameric constructs revealed no homology to the known plant peroxisome targeting signals PTS1, PTS2 or PEX19BS that are involved in transporting proteins to the peroxisomes[188–191]. Notably the constructs containing mCherry also often show fluorescent signal in the vacuole, with the majority of the signal present there by day 7 of protein expression (Figure 4.2T-U).

Instead of forming a large ER-derived compartment, these constructs followed a different pathway. After expression in the ER, the proteins are likely mistargeted for export to the peroxisomes. This process could be a detoxification mechanism aimed at removing excess or misfolded proteins from the ER. An alternative explanation is that these proteins accidentally hijack the peroxisome biogenesis pathway which originates in the ER[192]. The presence of mCherry from these constructs in the vacuoles suggests that further downstream of the peroxisomes, the constructs are eventually transported to the vacuole, possibly as a subsequent step in the detoxification mechanism or as a natural endpoint to the peroxisome lifecycle. mCherry fluorescence is more resistant to low pH than GFP, explaining why the GFP signal was not observed in the acidic vacuole[193], but it is likely that the GFP containing constructs follow the same ER – peroxisome - vacuole pathway. It is also possible that the constructs independently enter both the peroxisome and the vacuole directly from the ER, and an ER – peroxisome and ER – vacuole pathway are both present.

A notable shared property of most of these constructs is that their cytosolic oligomerisation domain is missing or oligomerises in a parallel fashion, supporting the hypothesis that a cytosolic facing antiparallel oligomerising domain is essential for the correct formation of the ER-derived compartment. The exceptions to this are the two

tetrameric constructs where the cytosolic domain is NHR2 (mC-22-NHR2 and G22NHR2), an antiparallelly binding tetrameric coiled coil. However, NHR2 tetramers are formed from two parallel NHR2 dimers[173]. One explanation to this discrepancy is that NHR2 tetramers are unable to form on the cytosolic side of the ER, potentially due to spatial hindrances, and only the parallel NHR2 dimers can form, which are insufficient to trigger the formation of the compartments and instead promote the hypothesised ER – peroxisome – vacuole pathway. However, further investigation is required to fully resolve this and to understand this branchpoint in the process of the compartment formation.

4.3.3 The effect of the protein construct's modular domains on the morphology of the compartment

The second-generation constructs investigated here bring the total number of G22Y derived constructs investigated in thesis or previously by Ian Moore's group to 30[155,156]. With such a large number of constructs, it is possible to categorize them based on compartment morphology and attempt to find common patterns, to lead to an overall hypothesis as to how the ER-derived compartment formation is affected by the modular domains of the polyproteins.

From the 30 constructs, six distinct phenotypes have been observed (Table 4.2). When no transmembrane domain is present in the polyprotein, the construct enters the cytosol and forms no obvious aggregate or compartment ("P1"). When the cytosolic oligomerisation domain is removed or is binding in a parallel fashion, the proteins mistarget to the hypothesised ER – peroxisome – vacuole pathway ("P2"). As suggested in Chapter 3.3.1, it is likely that parallel binding of oligomerising domains restricts the oligomerisation of the proteins spatially, only allowing two polyproteins from the same membrane to interact,

whereas antiparallelly binding oligomers can anchor together separate membranes, which could be an essential requirement for the formation of the ER-derived compartment.

When an antiparallelly binding cytosolic oligomerisation domain was present, four different compartment morphologies have been described (Table 4.2). The original G22Y protein and others where both the cytosolic and ER lumenal oligomerisation domains are antiparallelly binding show helical patterning (“P3”), likely caused by a repeating angle in the membrane curvature, induced by repeating oligomerisation events on either side of the membranes holding the scaffold proteins. The removal of the lumenal oligomerisation domain in 22Y likely removes this repeating angle in the membranes and generates a compartment with long fibrous parallel strands (“P4”). The replacement of the lumenal oligomerisation domain with a parallelly binding one (e.g., C22Y) leads to the formation of compartments with densely packed membranes which can’t be resolved by airy-scanning fluorescent microscopy (“P5”). This suggests, that while the antiparallel cytosolic oligomerising domain is essential for the formation of the compartments, the lumenal oligomerising domain also has an effect on the compartment morphology, depending on its binding orientation.

The addition of a C-terminal SpyTag and/or 6xHisTag lead to the formation of the sixth compartment structure (“P6”), similar to P5, but with notable small spherical holes in the fluorescent signal (Table 4.2). These holes are of the right dimensions to be organelles (e.g., peroxisomes, Golgi, or mitochondria, but not chloroplasts, since those would have been identified due to their autofluorescence) or other native subcellular compartments embedded into the large synthetic compartments generated by my constructs. The identification of these organelles could be carried out by co-expressing the compartment

forming constructs with a range of appropriate fluorescent markers for the likely organelle candidates.

The same morphology was observed for a variety of constructs, despite the fact that these constructs featured a range of binding domains of different size, binding affinity and binding orientation (on the lumenal side). Likewise, removal or lengthening of the linkers between the cytosolic binding domain and the SpyTag yielded no changes to the compartment. Thus, since the appearance of this morphology is closely linked to the addition of a C-terminal SpyTag and/or 6xHisTag (in fact, addition of either or both of these tags to compartment forming constructs almost exclusively leads to the formation of this phenotype) it is possible that these tags are responsible for the capture of the non-fluorescent native compartments via a charged-based interaction. However, the tags are not sufficient on their own to produce the compartments, since the removal of the cytosolic oligomerisation domain on the construct mC-C22-SpyT-his lead to the proteins following the hypothesised ER – peroxisome – vacuole pathway and not producing a compartment.

Notably there is one other exception to the rule that a C-terminal tag always causes the formation of this morphology: Cap22Y-SpyT-his unexpectedly forms helical compartments similar to G22Y (P3) instead of the expected P6. While the oligomerisation domains of both compartments are all antiparallel, it is unclear why the C-terminal tags did not cause the formation of the expected compartment in this case and why the construct “defaulted” back to the compartment structure expected based on the oligomerisation domains alone.

It is important to note, that for most second-generation constructs, the addition of the C-terminal tags was carried out in parallel to other modifications, since it was initially

assumed that the C-terminal tags will not have a major effect on the compartment morphology. As such, the effect of the C-terminal tags confounded the effect of other domain changes. It is thus recommended for future work to reduce changes to the compartment-forming protein scaffolding to a single change at a time to eliminate this confounding factor.

Compartment morphology	Constructs producing this morphology	Common denominator of constructs
Diffuse non-ER signal (no compartment) – "P1"	G-Y(cyt)	No transmembrane domain
ER - peroxisome - vacuole mistargeting (no compartment) – "P2"	G22C, G22, G22mRFP*, mC-C22-SpyT-his, G22-NHR2, G22-CCTet, mC-22-NHR2, mC-22-CCTet, NHR2-22-mC, CCTet-22-mC	Cytosolic oligomerising domain is missing or is parallelly binding
Helical strands forming large compartment – "P3"	G22Y, NHR2-22Y, Cap22Y-SpyT-his	Both oligomerising domains are antiparallelly binding
Fibrous structures forming large compartment – "P4"	22Y	Only a single, cytosolic antiparallelly dimerising domain
Large compartments with even fluorescent signal coverage of surface – "P5"	C22Y, mRFP-G22Y*, mRFP-22Y*, CCTet-22Y	Cytosolic oligomerisation domain is antiparallelly binding and lumenal oligomerisation domain is parallelly binding
Large compartments with spherical holes in fluorescent signal – "P6"	C22Y-SpyT-his, C22Y-ΔSpyT-his, C22Y-SpyT-Δhis, C22Y-SL-SpyT, C22Y-NL-SpyT, 22LLY-SpyT-his, G22Cap-SpyT-his, mC-C22FDH-SpyT-his, mC-22FDH-SpyT-his, mC-C22FDH-ΔSpyT-his, mC-C22FDH-SpyT-Δhis	Cytosolic oligomerisation domain is antiparallelly binding and at least one C-terminal tag is present

Table 4.2: Categorisation of compartment morphologies and constructs. Constructs marked with an asterisk were described previously in previous works produced by the Moore lab[155,156]. Phenotypes were given a number (P1-P6) for ease of recognition in the discussion.

Based on the patterns described, the framework for a testable general hypothesis on the driving forces of the ER-derived self-organising compartment formation can be assembled (Figure 4.8). Three essential checkpoints appear to exist before a compartment can form. First, an ER-targeted transmembrane domain is required, for correct localisation of the protein scaffold and ER-membrane recruitment, which is essential to compartment formation. Second, the polyprotein expression levels need to be sufficiently high to trigger compartment formation. It was a general observation for all constructs that when expressing at very low levels, no compartment forms, and the proteins remain in the unmodified ER. Third, a cytosolic, antiparallelly binding oligomerising domain is required as part of the polyproteins to avoid the ER – peroxisome – vacuole pathway. It is likely that the formation of the compartment protects the proteins from this fate, but in absence of this domain, the compartments cannot form.

If all three checkpoints have been passed, the minimum requirements for compartment formation have been met, but the final compartment structure is dependent on the orientation of the oligomerising domains and the addition of any C-terminal tags, as described previously. While the current suite of constructs almost exclusively follows this hypothesis (with the exceptions of Cap22Y-SpyT-his as previously noted), it cannot be excluded that more compartments with novel morphologies could be generated and additional rules elucidated on the formation of the compartments with the design and development of new constructs, opening up exiting possibilities for further research in this topic. The general applicability of these rules is reinforced by the fact that constructs with higher-order coiled coils broadly followed them as well.

A final question which I can attempt to answer using these results is whether the compartments described here are OSER or indeed a completely novel class of ER-derived compartment. The checkpoints for the compartment formation are remarkably similar to those of the prevalent OSER formation models, most notably the 'unified theory of OSER formation' (Figure 1.3) where both sufficiently high membrane protein expression levels and antiparallel oligomerisation domains have been suggested as triggers for OSER formation[143]. The compartment-forming protein scaffold structure resembles that of canonical OSER-inducing constructs (i.e.: an ER-resident transmembrane domain with a cytosolic head-to-head oligomerising domain fused to it) and there is no apparent incompatibility between the hypothesis of the compartment formation checkpoints (Figure 4.8) and the 'unified theory of OSER formation'. While the compartments can be morphologically distinct from many plant OSERs in the literature[132,144,194], similar OSER structures have been described as well[148,151,181]. It must also be noted that OSER itself is a very broadly defined category with no universally agreed classification method[143], and as such, the compartments described in this thesis could be plausibly categorised as a subgroup of OSER. Indeed, with the growing literature on OSER, it might be beneficial to establish a more standardised classification system. This should combine the 'unified theory of OSER formation' and the morphological observations of the formed structures to provide information on both the mechanism of formation and the final ultrastructure. Effective categorisation of OSER phenotypes could shed more light on their common properties, formation and their potential functionalisations as a group.

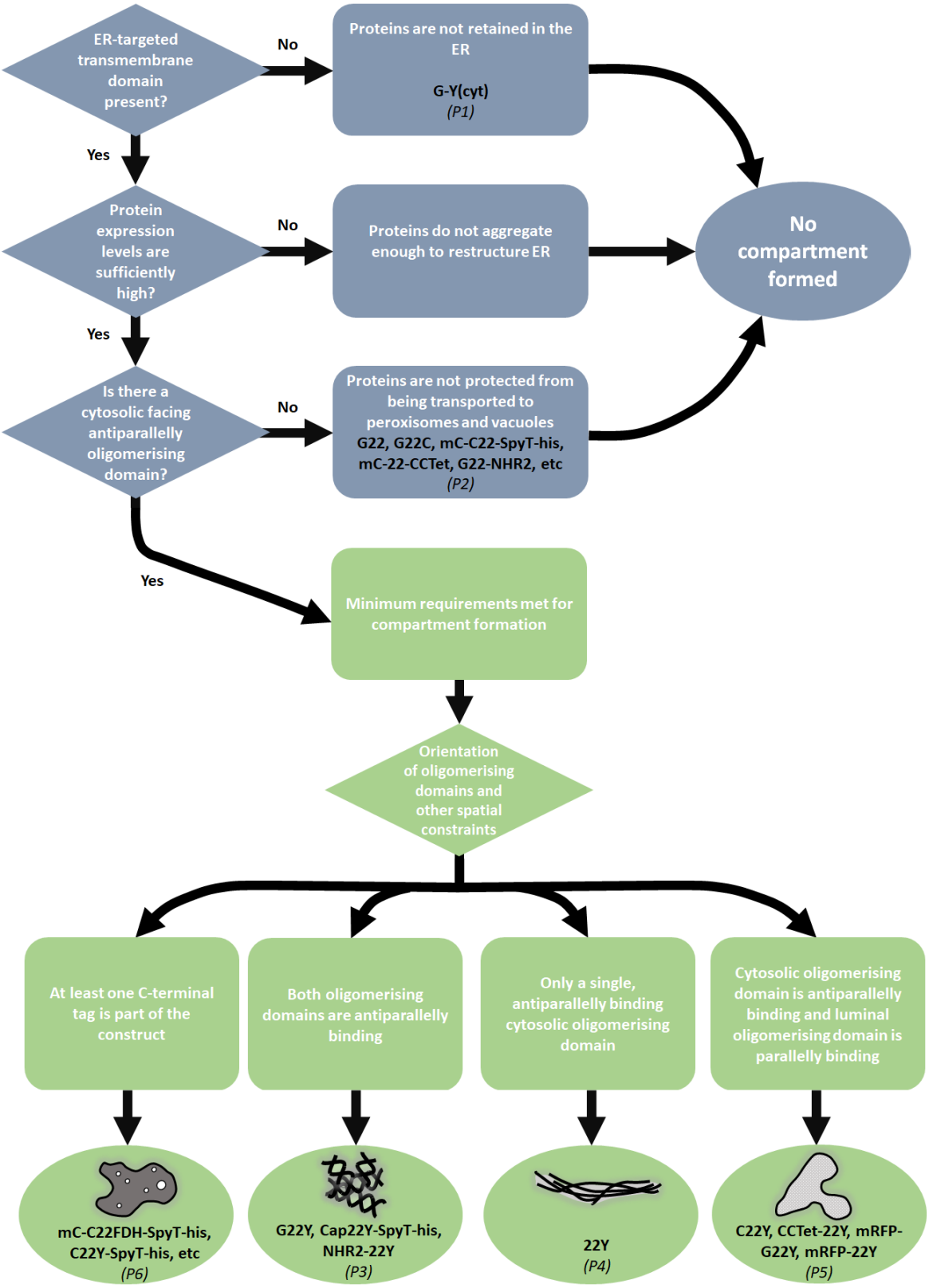

ER-targeted transmembrane domain present?
No
Proteins are not retained in the ER
G-Y(cyt)
(P1)
Yes
Protein expression levels are sufficiently high?
No
Proteins do not aggregate enough to restructure ER
No compartment formed
Yes
Is there a cytosolic facing antiparallelly oligomerising domain?
No
Proteins are not protected from being transported to peroxisomes and vacuoles
G22, G22C, mC-C22-SpyT-his, mC-22-CCTet, G22-NHR2, etc
(P2)
Yes
Minimum requirements met for compartment formation
Orientation of oligomerising domains and other spatial constraints
At least one C-terminal tag is part of the construct
Both oligomerising domains are antiparallelly binding
Only a single, antiparallelly binding cytosolic oligomerising domain
Cytosolic oligomerising domain is antiparallelly binding and luminal oligomerising domain is parallelly binding
mC-C22FDH-SpyT-his, C22Y-SpyT-his, etc
(P6)
G22Y, Cap22Y-SpyT-his, NHR2-22Y
(P3)
22Y
(P4)
C22Y, CCTet-22Y, mRFP-G22Y, mRFP-22Y
(P5)

Figure 4.8 (above): The general hypothesis on the driving forces and checkpoints for the ER-derived compartment formation. The three major checkpoints in grey (presence of ER-targeted transmembrane domain, sufficiently high protein expression levels, cytosolic facing antiparallelly binding oligomerisation domain) are essential for compartment formation and lack of either will lead to no large compartment forming. Once these checkpoints have been passed, and the minimum requirements for compartment formation have been met (in green), the spatial orientation of oligomerising domains and presence of other domains can have varying effects on compartment structure, leading to a range of possible compartments. Construct names in black show examples of constructs belonging to the given category. Italics show the compartment phenotype categories observed.

Chapter 5: Foundation steps towards application

5.1 Introduction

A group of large synthetic membranous ER-derived compartments whose formation is induced by the expression of a single protein construct was discovered and characterised in Chapter 3. Many of the compartments' properties made them an attractive potential target for functionalisation. The compartment-forming scaffold proteins were modular, and it was easy to genetically fuse novel protein tags to them. The compartments had little effect on the rest of the ER and showed an effective capacity to provide a diffusion barrier to the cytosol. As such, the development of this system as a platform to produce difficult-to-express or toxic membrane proteins or as a metabolic microdomain for probabilistic metabolic channelling was immediately raised as a possibility. With the realisation in Chapter 4 that these compartment types are likely a subgroup of OSER, the drive for functionalisation was increased, since such a possibility has been raised before in the literature, but never realised to its full potential[31,143].

The basis of functionalisation was the addition of two new peptide tags in Chapter 4: a 6xHisTag for efficient and quick protein purification and a SpyTag, which can form a covalent bond with the corresponding SpyCatcher[184,185], allowing the modular anchoring of further cargo proteins to the cytosolic surface of the compartment. These new scaffold constructs were characterised to ascertain that even with the addition of these tags, the compartments retain the properties that made them candidates for functionalisation in Chapter 3.

As they did, a series of proof-of-principle experiments still had to be conducted to begin the testing of the compartments as a synthetic biology tool. In this chapter, I describe these

pilot studies. I have tested the purification of the constructs using both nickel-affinity purification enabled by the addition of the 6xHisTag and an immunoprecipitation strategy. I have also tested the SpyCatcher – SpyTag system to covalently attach cytosolic cargo proteins to the surface of the compartment and then purify them together. The attachment of cargo proteins was also used to determine if enzymes anchored to the cytosolic surface of the compartments remain functional in their new subcellular microenvironment. An experiment to determine if the compartment is an effective microdomain was proposed and prepared, but was unsuccessful and due to time-constraints could not be troubleshooted.

Finally, I characterised *A. thaliana* lines that stably express the compartment-forming proteins. This allowed me to assess if the plants remained viable when they express the compartments throughout their life cycle and in all of their tissue types which would facilitate the future use of the compartments in a variety of biotechnology applications.

5.2 Results

5.2.1 Developing a method to purify the compartment-forming protein constructs

A reliable method to purify the compartment-forming protein constructs was required as a first step towards recovery of valuable proteins genetically fused or post-translationally covalently attached to the compartment. To facilitate this and enable nickel-affinity purification, 6xHisTags were genetically fused to several of the constructs described in Chapter 4.

Prior to purification, it was important to measure the proportion of total cellular fluorescent signal present in the compartment. Having a very high proportion of the compartment-forming constructs inside the compartment was essential, since the applications proposed for the compartments all relied on the effective concentration of the target proteins and minimal scaffold leakage into the rest of the cell. To this end, whole-cell Z-stacks were captured using laser-scanning fluorescence confocal microscopy on mature *N. benthamiana* leaves 7 d after *A. tumefaciens*-mediated transient transformation with either C22Y or C22Y-SpyT-his. *N. benthamiana* was used for experiments in this chapter due to its improved amenability to higher level expression of transiently transformed constructs. The fluorescence signal inside the compartments was divided by total cellular florescence (Figure 5.1A). A high proportion of the total cellular fluorescence was observed to be restricted to the compartment for both scaffold types, with C22Y (mean = 92.7%) outperforming C22Y-SpyT-his (mean = 82.8%), showing a significant statistical difference ($p = 0.0039$) (Figure 5.1B). However, since even C22Y-SpyT-his had

most of its fluorescence inside the compartments, it was deemed suitable for downstream applications.

C22Y-SpyT-his was purified from transiently transformed mature *N. benthamiana* leaves using nickel-affinity purification via gravity flow chromatography. Samples were taken at every step of the process and the sample fluorescence quantified to follow the recovery of the fluorescent compartment-forming protein construct (Figure 5.1C). The polyhistidine tags proved to be efficient for one-step purification, as fluorescence followed the expected pattern throughout the purification steps for all constructs: untransformed leaves showed no fluorescence, while extracts of C22Y expressing leaves had substantial fluorescence that washed straight through the nickel-NTA resin. In contrast, purification of C22Y-SpyT-his containing extracts resulted in considerable fluorescent signal interacting with the resin that could be specifically eluted (Figure 5.1C). When the fluorescence was quantified, recovery rates were estimated to be between 22-56%. Purification was confirmed by anti-GFP Western Blots of the elution fractions which showed bands with the appropriate molecular weight for C22Y-SpyT (Figure 5.1D), but no bands for any other samples.

Noticeably, some faint higher molecular weight bands also appeared, which were approximately the right size for C22Y-SpyT-his dimers despite the fact that dimerization was not expected to be possible due to the addition of SDS and the harshly reducing DTT. However, when the SDS-PAGE gels were directly fluorescently visualised, both the lighter and heavier bands detected using Western Blot showed strong YFP fluorescence, suggesting that the YFP molecules retained their structure, and thus possibly their capacity for dimerization (Figure 5.1E). In-gel fluorescence of GFP-tagged membrane proteins have

been described before, in some cases including dimerization as well[195–197], possibly due to the YFP's ß-barrel structure resisting denaturation.

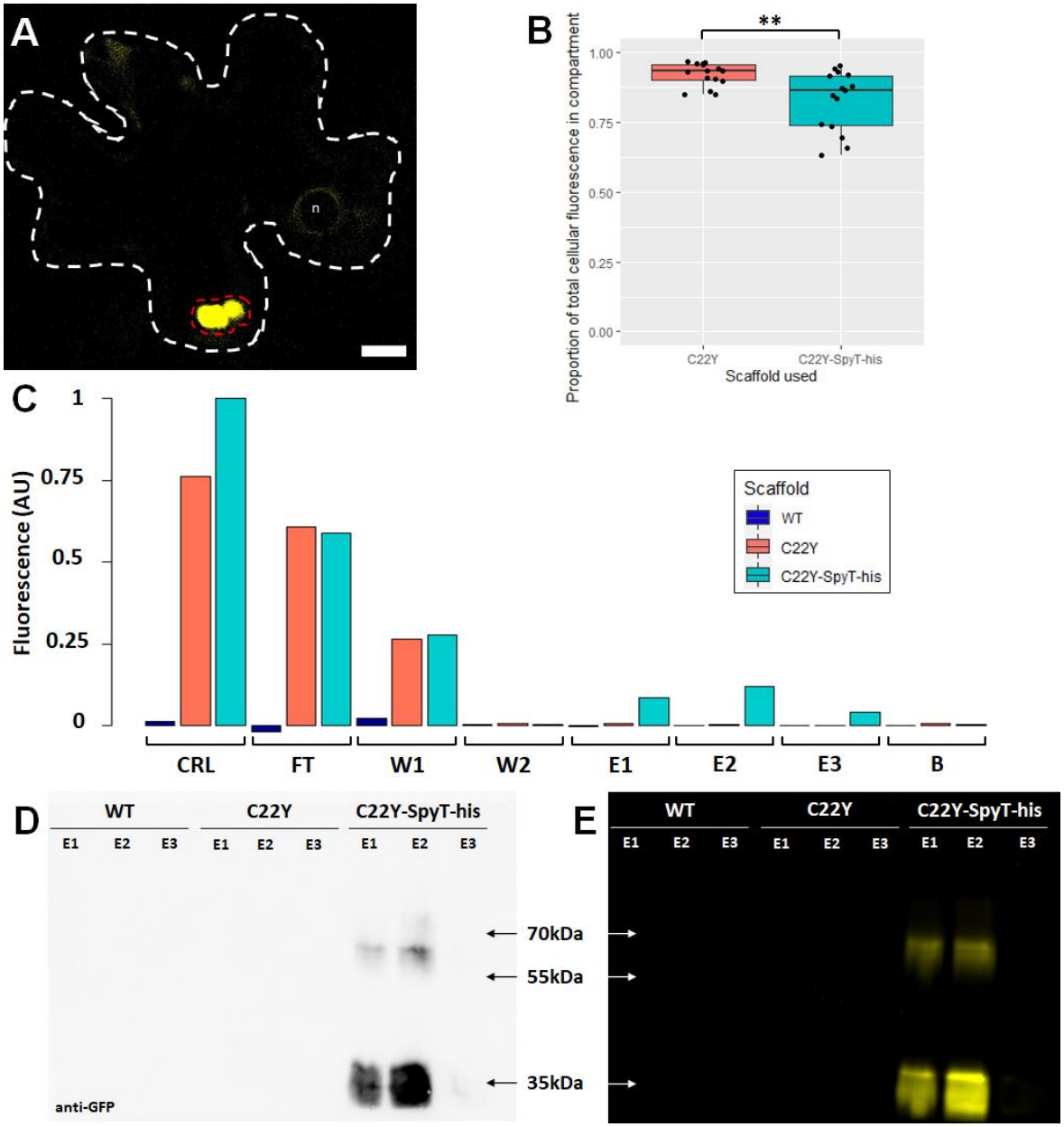


Figure 5.1: Purification of the compartment-forming scaffold proteins. A: Example fluorescent confocal microscopy image from a full-cell Z-stack of mature *N. benthamiana* leaves expressing C22Y-SpyT-his 7 d after *A. tumefaciens*-mediated transient transformation. The cell is outlined with white dashed lines, and the compartment with red dashed lines, showing the boundaries used when calculating total cellular fluorescence and compartment fluorescence. The image here is brightened for clarity but was not saturated for measurement purposes. Scale bar: 10 µm. B: The proportion of the total cellular fluorescence presence in the compartments. Both C22Y and C22Y-SpyT-his have a high retention rate of fluorescent molecules inside the compartment. The statistical

difference is significant (p = 0.0039) between the two scaffold types. The boxes represent the interquartile range, the horizontal line in the box shows the median, and the whiskers the minimum and maximum values (excluding outliers). n = 15 for all categories. C: Typical fluorescence (AU – normalised to maximum detected value per µL sample) of each fraction during nickel-affinity purification via gravity flow chromatography. CRL: crude lysate; FT: flowthrough after binding; W1: first wash; W2: last wash; E1-E3: successive elution fractions; B: resin beads post-elution. WT samples show no fluorescence. C22Y show high fluorescence but as it has no his-tag, this is lost during flow-through and washes. C22Y-SpyT-his can be recovered during elution steps. D: Anti-GFP Western Blot of elution fractions (E1-E3) from C. C22Y-SpyT-his bands of expected molecular weight (40 kDa) can be seen. A faint heavier band, approximately the expected molecular weight of a C22Y-SpyT-his dimer can be also seen. E: In-gel YFP fluorescence of the SDS-PAGE gel used to prepare D. Both the lighter and heavier C22Y-SpyT-his bands clearly fluoresce.

5.2.2 Proof-of-principle purification of a genetically fused, compartment-localised enzyme

A simple way of attaching a protein of interest to the compartment is genetic fusion to the compartment-forming construct. This was tested using a proof-of-principle purification of the mC-C22FDH-SpyT-his construct described in Chapter 4, which uses a formate dehydrogenase enzyme as a cytosolic dimerizing domain. A novel protein construct was generated to serve as a suitable cytosolic control for this purification experiment, with the ER luminal facing CC-Di coiled-coil domain, the BP22 transmembrane domain and the ER-targeting signal peptide removed to produce mC-FDH-SpyT-his (Figure 5.2A). Unlike mC-C22FDH-SpyT-his, which forms large compartments (Figure 4.4I & 5.2B), mC-FDH-SpyT-his remains mostly in the cytosol, with some signal in the vacuole, without forming any compartments (Figure 5.2C).

To diversify the portfolio of available tools for purification, an mCherry immunoprecipitation method was tested on these constructs. This method also gave

opportunity to test if a luminally targeted purification method can be successful. Anti-mCherry Western Blot analysis of the elution fraction showed bands of the expected size for both constructs (Figure 5.2D), confirming successful purification via an N-terminal tag of an enzyme that was genetically fused to the cytosolic C-terminus of the transmembrane domain. This opens up the possibility of using the compartment as a protein expression tool for difficult-to-express targets, such as membrane proteins. Further research into the use of N-terminal (luminal) tags is recommended as these could also open up further functionalisations (i.e., purification of protein that are captured before leaving the ER).

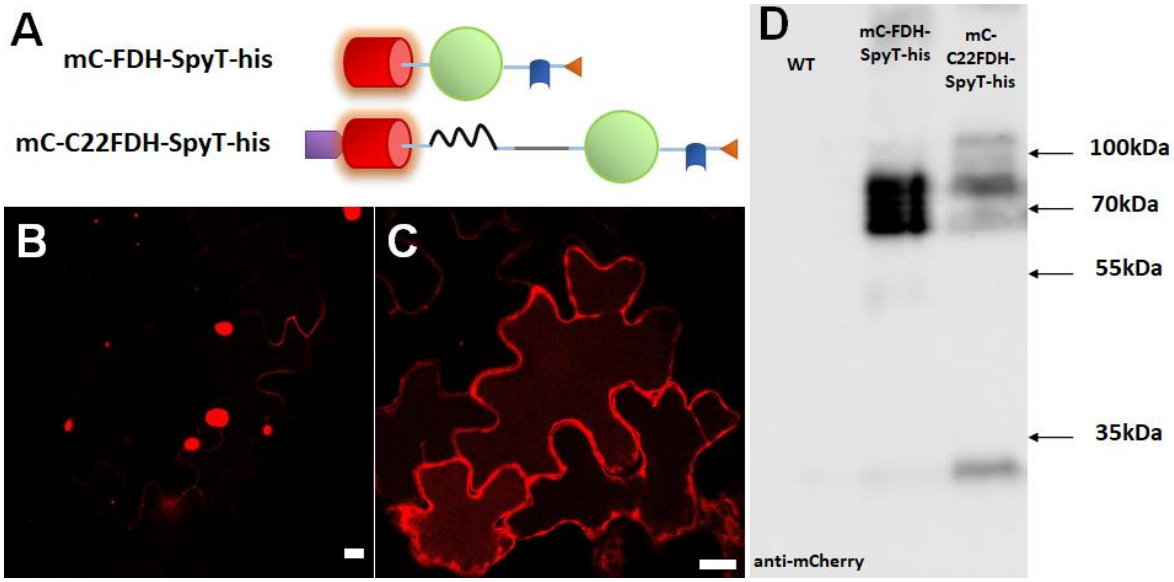


Figure 5.2: Purification of a genetically fused enzyme via immunoprecipitation. A: The two constructs used for this experiment. The new control construct mC-FDH-SpyT-his has had its ER-targeting signal, ER-luminal CC-Di coiled-coil domain and BP22 transmembrane domain removed. B & C: Fluorescent confocal microscopy images (red: mCherry fluorescence) of mC-C22FDH-SpyT-his (B) and mC-FDH-SpyT-his (C) in mature *N. benthamiana* leaves 7 d after *A. tumefaciens*-mediated transient transformation. As expected, the truncated control construct had its compartment forming capacity abolished. Images were taken at different settings and are not appropriate for direct comparison of fluorescence levels. Scale bars: 10 µm. D: Anti-mCherry

Western Blot of samples purified via immunoprecipitation. Both mC-FDH-SpyT-his and mC-C22FDH-SpyT-his show bands of expected size (73 kDa and 84 kDa, respectively). mC-C22FDH-SpyT also has a band with molecular weight approximately 30 kDa, which fits the expected molecular weight of a cleaved mCherry protein.

5.2.3 Assessing the efficiency of the compartment as a custom surface area for anchoring proteins

Beside genetic fusion of target proteins to the compartment-forming construct, the other option to attach them to the surface of the compartment is post-translationally via covalent binding driven by the SpyCatcher – SpyTag system[184,185]. To this end, the C22Y-SpyT-his construct was developed and characterised in Chapter 4. This construct enables a modular approach, as any cargo protein can be bound to the cytosolic surface of the compartment which has the appropriate SpyCatcher peptide. The cargo protein template which was used for all cargo constructs in this Chapter also contains a 6xHisTag for nickel-affinity purification, and an mCherry as a fluorescent protein tag (Figure 5.3A).

Due to the presence of the mCherry in the cargo, the binding of the cargo to the compartment can be tested using laser-scanning fluorescence confocal microscopy. The normally cytosolic cargo construct should stay in the cytosol when co-expressed with a C22Y construct without the SpyTag, and the mCherry fluorescence would not be expected to coincide with the YFP from C22Y. However, when co-expressing a SpyCatcher containing cytosolic cargo with a C22Y-SpyT-his construct, most of the mCherry signal from the cytosol is expected to be associated with the compartment, overlapping with the YFP signal (Figure 5.3B).

Two cargo constructs were utilised to confirm covalent binding to the compartment via the SpyCatcher – SpyTag system. The first contained the luciferase enzyme (Luc) from the

firefly *P. pyralis*, while the second was the 2-hydroxybiphenyl-3-monooxygenase (HbpA) from the 2-2'-biphenol breakdown pathway of *P. nitroreducens*. Both were co-expressed with C22Y and C22Y-SpyT-his (Figure 5.4). Luc was co-expressed with the compartment-forming constructs simultaneously, while HbpA was expressed sequentially: transformation with the compartment-forming constructs was followed by the introduction of HbpA 7 d later, after allowing time for the compartments to form.

Both Luc and HbpA showed a clear pattern of co-localising with C22Y-SpyT-his compartments but appeared throughout the entire cytosol when co-expressing with C22Y, as expected (Figure 5.4). This suggested that compartment formation is not inhibited by the simultaneous co-expression and bindings of a cargo protein (e.g., by spatial interference of the YFP dimerization on the cytosolic side). Furthermore, the sequential co-expression also showcased that the centre of fully formed compartments is accessible to cargo proteins (Figure 5.4).

To assess the efficiency of the cargo capture, a series of whole-cell Z-stack laser-scanning fluorescence confocal microscopy images were taken, and the proportion of total cellular mCherry fluorescence in the compartment was measured (as shown in Figure 5.1A). For both Luc and HbpA the difference was statistically significant when comparing co-expression with C22Y (mean for Luc = 7.3%; mean for HbpA = 12.0%) and C22Y-SpyT-his (mean for Luc = 48.9%, $p = 4.7 \times 10^{-7}$; mean for HbpA = 87.0%, $p = 1.3 \times 10^{-16}$), in line with the visual inspection (Figure 5.5).

HbpA had a clearly greater proportion co-localising with the C22Y-SpyT-his compartment than Luc, but quantification of this would be not appropriate since it is unclear if the difference is caused by the different expression strategies or the expression levels of the

cargos. However, the results suggest that sequential co-expression could possibly be a better strategy to improve capture ratio of cargo proteins. Intuitively this could be explained by the fact that compartments require time to form, during that period the cargo proteins cannot bind to them and might be broken down or transported out of the cytosol and be inaccessible to the surface of the compartment when it later forms. This period is abolished when expressing the two constructs sequentially, as any translated cargo proteins will immediately have access to a fully formed compartment. An alternative explanation could be simply much higher expression of Luc than C22Y-SpyT-his, and saturation of the compartment surface. However, this is unlikely, since the compartment forming constructs generally expressed to a higher level than the cargos, and HbpA usually had a greater fluorescence level than Luc. Further experiments comparing the efficiency of sequential transformation and simultaneous transformation (using the same cargos) are recommended to clarify this.

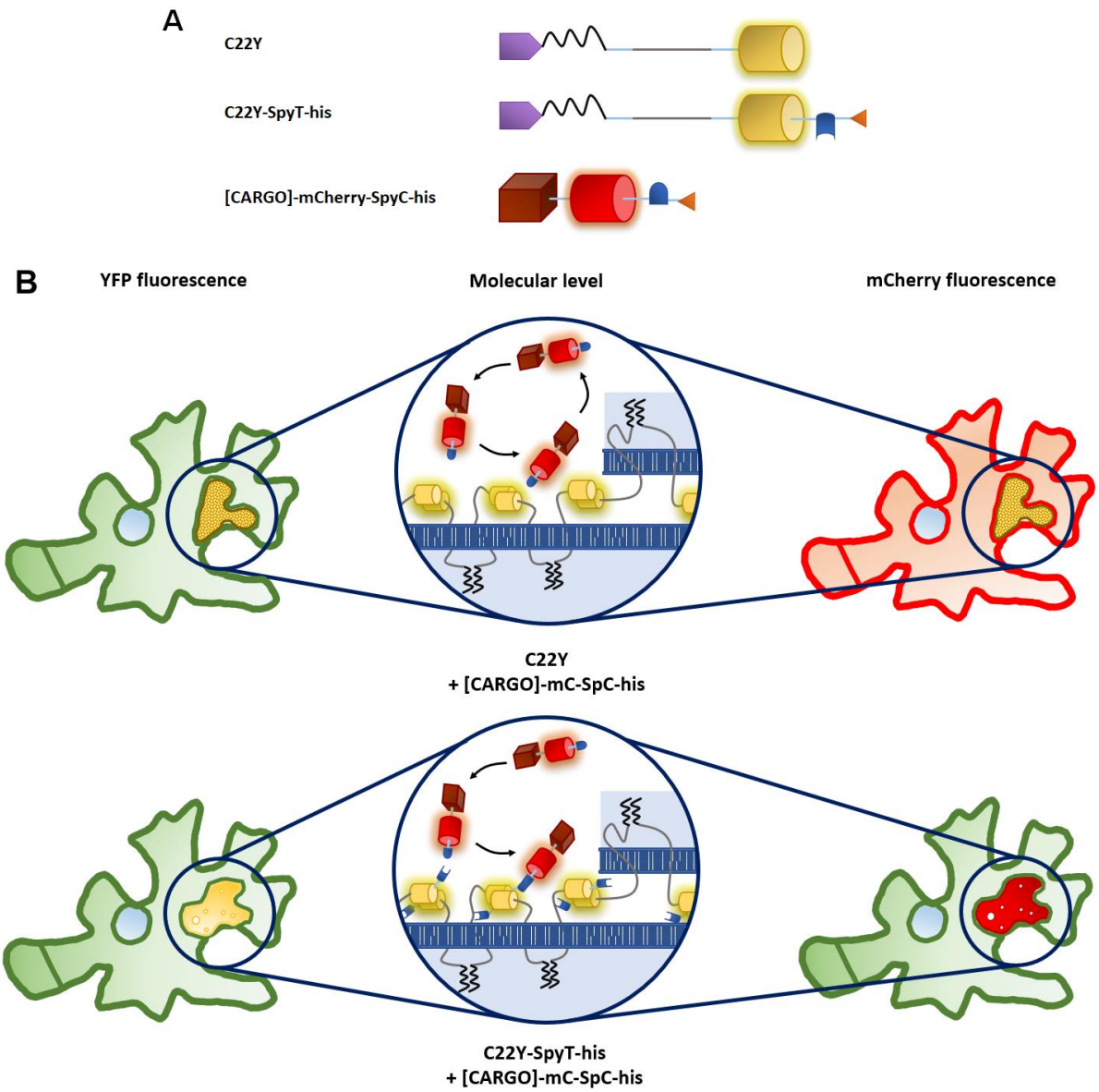


Figure 5.3: Schematic representation of SpyCatcher – SpyTag binding to the surface of the compartment. A: Two scaffold construct were used, one with SpyCatcher and a 6xHisTag (C22Y-SpyT-his) and one without (C22Y). The cargo construct can be modular, with any protein genetically fused to a cytosolic mCherry, SpyCatcher and a 6xHisTag – all cargo constructs used in this chapter use this template. B: When co-expressing a cargo with a compartment construct, if the compartment has no SpyTag, covalent binding is not possible and the YFP signal from the compartment and the mCherry signal from the cargo is not expected to overlap. When the compartment has a SpyTag peptide, most of the mCherry signal is expected to be concentrated into the compartment due to the covalent binding of the cargo to the C22Y-SpyT-his scaffolds.

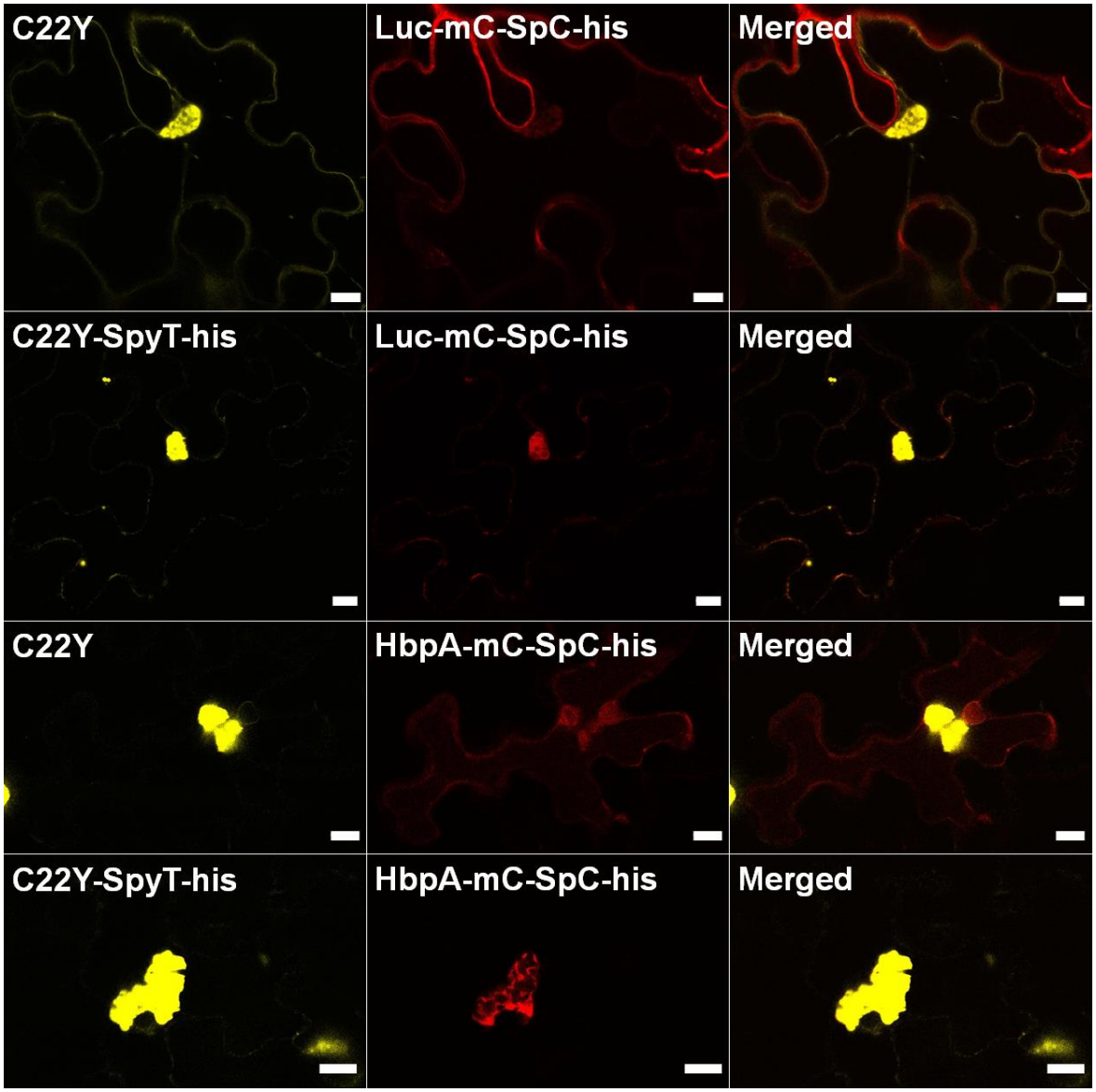


Figure 5.4: Visualisation of proteins anchored to the cytosolic surface of the compartment via SpyCatcher – SpyTag covalent binding. Fluorescent confocal microscopy images (red: mCherry fluorescence, yellow: YFP fluorescence) of HbpA-mC-SpC-his (HbpA) and Luc-mC-SpC-his (Luc) with C22Y/C22Y-SpyT-his in mature *N. benthamiana* leaves 7 d after *A. tumefaciens*-mediated transient transformation. Note that for the HbpA images a sequential transformation protocol was used, with the C22Y/C22Y-SpyT-his constructs introduced 14 d before the images were taken. Scale bars: 10 µm.

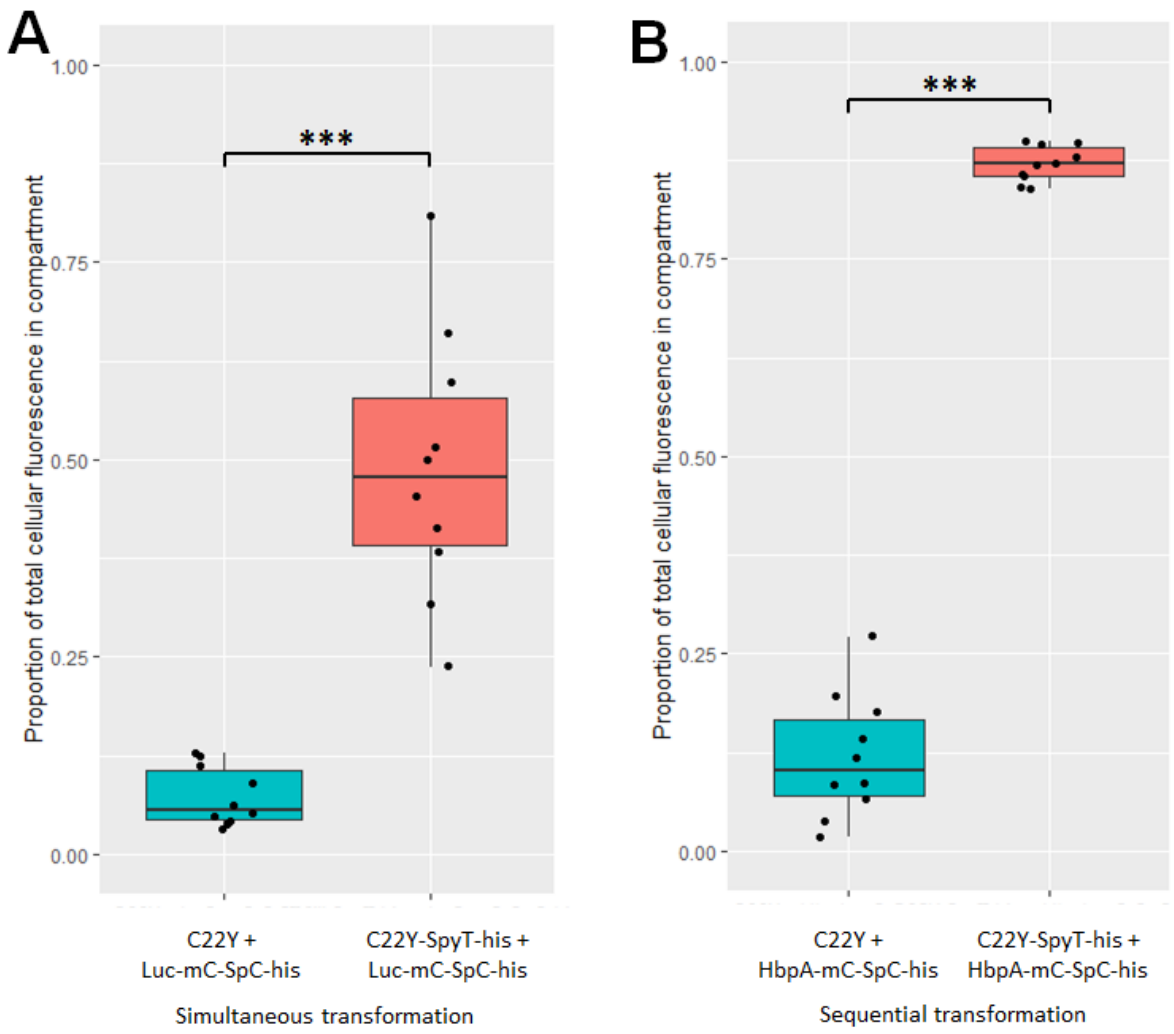


Figure 5.5: Efficient recruitment of proteins to the cytosolic surface of the compartment via SpyCatcher – SpyTag covalent binding. The proportion of total cargo (Luc (A) or HbpA (B)) mCherry fluorescence in the compartment. Both cargo types show a statistically significant difference between their co-expression with C22Y and C22Y-SpyT-his (for Luc $p = 4.7 \times 10^{-7}$; for HbpA $p = 1.3 \times 10^{-16}$). The boxes represent the interquartile range, the horizontal line in the box shows the median, and the whiskers the minimum and maximum values (excluding outliers). n = 10 for all categories

5.2.4 Purification of enzymes anchored to the surface of the compartment

Given that I was able to demonstrate successful attachment of proteins to the surface of the compartment via the SpyCatcher – SpyTag system the next step was to attempt the purification of these cargo-carrying complexes. Since both the compartment-forming constructs and the cargo proteins had a 6xHisTag and nickel-affinity purification via gravity flow chromatography was successfully used with C22Y-SpyT-his before (Figure 5.1), the same method was used on mature *N. benthamiana* leaves with sequential transient transformation using two cargo types: the previously used HbpA construct, and an extradiol ring cleavage dioxygenase (HbpC) from the 2-2'-biphenol breakdown pathway of *P. nitroreducens.*

Both cargo constructs contained an mCherry protein as well, which was used to verify the purification via an anti-mCherry Western Blot (Figure 5.6). When co-expressing with a C22Y-SpyT-his construct, both cargos showed a new, heavier band in the elution fraction for the covalently bound cargo + scaffold molecules, and the band corresponding for the unbound cargo protein disappeared, which signalled a successful co-purification of the bound cargo with the scaffolding. The bands of HbpC consistently had an approximately 15 kDa lower molecular weight than expected, possibly due to protein cleavage or because some membrane proteins have the tendency to run quicker in SDS-PAGE gels than expected[198].

Notably, another set of lower molecular weight bands were consistently present in all samples, which did not increase in molecular weight when co-expressed with C22Y-SpyT-his. Since these bands are present in the elution fraction, they are expected to retain the C-terminal 6xHisTag, and identification via an anti-mCherry Western Blot suggests that the

mCherry was present too. Based on the molecular weights, these were hypothesised to be cleavage products with a part of the HbpA/C proteins lost. Curiously, despite having a SpyTag (since that is positioned between the 6xHisTag and the mCherry), neither band increased in molecular weight when co-expressed with C22Y-SpyT-his.

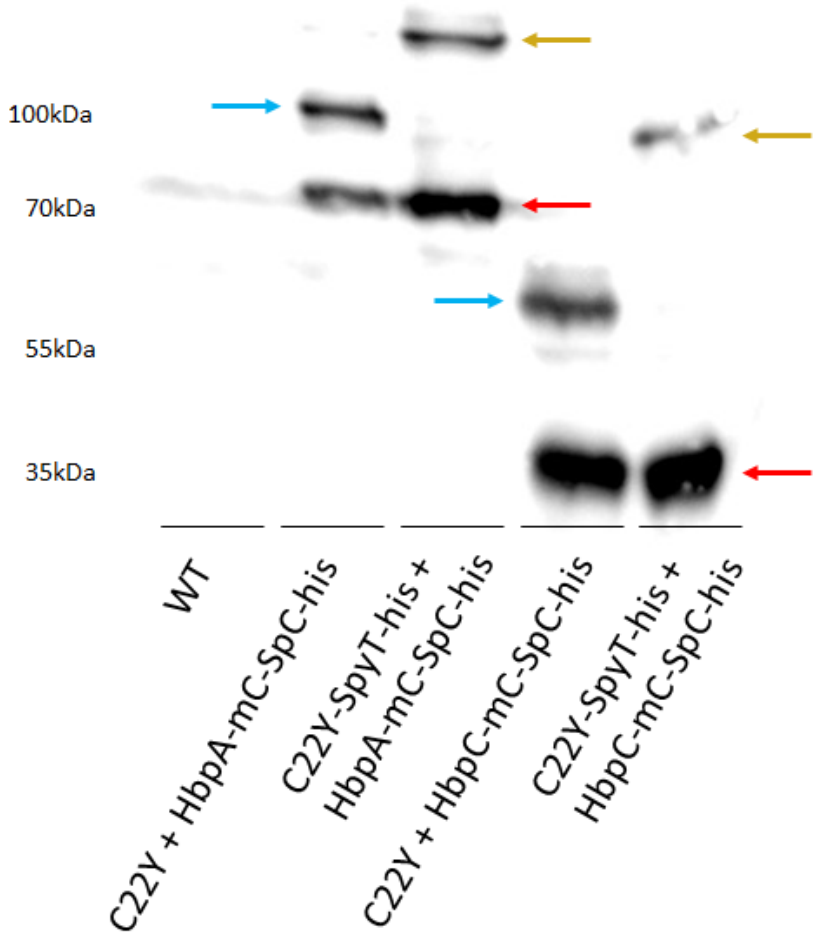


Figure 5.6: Purification of a covalently bound enzyme – protein-forming scaffold complex. Anti-mCherry Western Blot of elution fractions from a nickel-affinity purification from sequentially transformed *N. benthamiana* samples. Blue arrows show bands with matching molecular weight to unbound HbpA/HbpC-mC-SpC-his cargo constructs (107 kDa and 76 kDa, respectively). Golden arrows show bands with matching molecular weight to bound HbpA/HbpC-mC-SpC-his + C22Y-SpyT-his complexes (147 kDa and 116 kDa, respectively). Red bands show putative cleavage products. HbpC-mC-SpC-his samples consistently show approximately 15 kDa lighter bands than expected.

5.2.5 Testing the functionality of enzymes attached to the surface of the compartment

A simple way to determine if probabilistic channelling can be achieved in the compartment is to use a branching metabolic pathway where the intermediate metabolite is used by two competeing enzymes[83]. For such a pathway, if the enzyme producing the intermediate and one of the subsequent enzymes are both attached to the compartment and thus co-localised at high local concentrations, then flux through that branch should be higher than towards the competing pathway. This difference can be measured by comparing fluxes in a system with free enzymes to assess the efficiency of probabilistic metabolic channelling.

One metabolic pathway with appropriate branching is the 2,2'-biphenol breakdown pathway from *P. nitroreducens* (Figure 5.7). The enzyme HbpA turns 2,2'-biphenol into 2,2',3-trihydroxybiphenyl, which is released from the active site and then can be either continuously metabolised by HbpA into 2,2'3,3'-tetrahydroxybiphenyl or alternatively turned into (2Z,2E)-2-hydroxy-6-oxo-6-(2-hydroxyphenyl)-hexa-2,4-dienoate by the enzyme HbpC[199–201]. This pathway was selected for the testing of the compartment as a microdomain due to the low number of enzymes involved (and as such, the reduced complexity of the experiment), the cheap availability of the substrate, and the fact that none of the enzymes and metabolites are present in plants natively, so metabolic measurements are not affected by the host background. The enzymes were expressed and tested for attachment to the compartment (Figure 5.6) but enzymatic activity was never detected and no intermediate or final product was ever identified using GC-MS.

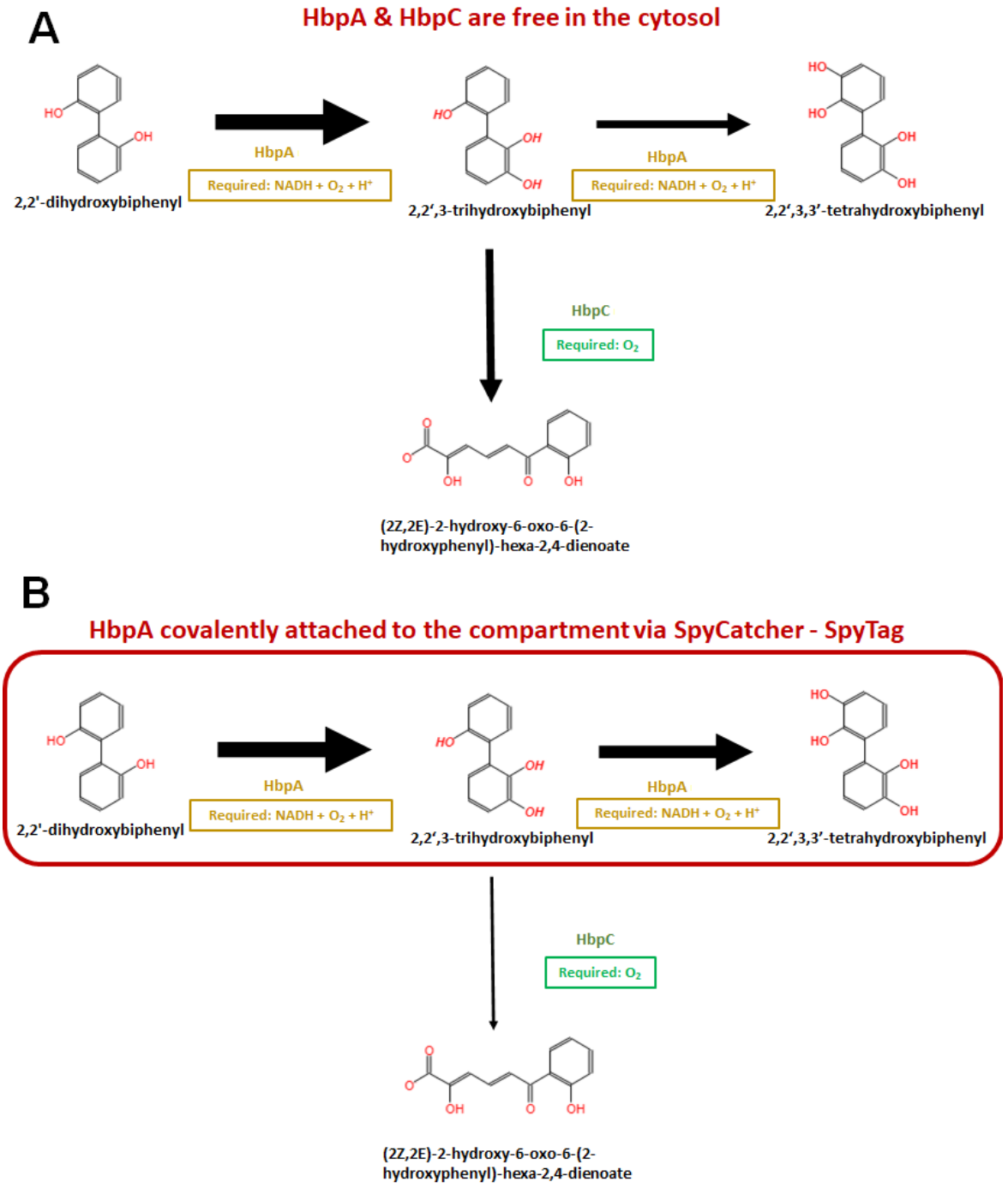


Figure 5.7: Testing the efficiency of a metabolic microdomain via the 2,2'-biphenol breakdown pathway. A: When using free (e.g., cytosolic) enzymes, the output of the two final products only depends on the kinetic properties and abundance of enzymes HbpA and HbpC which compete for the intermediate 2,2',3-trihydroxybiphenyl. When HbpA is co-localised at high concentrations to the compartment, the intermediate is more likely to encounter HbpA than HbpC (probabilistic metabolic channelling) increasing the concentration of 2,2',3,3'-tetrahydroxybiphenyl and decreasing (2Z,2E)-2-hydroxy-6-oxo-6-(2-hydroxyphenyl)-hexa-2,4-dienoate, which can be measured using GC-MS to determine the efficiency of metabolic channelling in the microdomain.

The failed microdomain experiment raised a potential issue: successfully attaching enzymes to the surface of the compartment is the first step in the creation of a metabolic microdomain, but the formation of a covalent bond is insufficient evidence on its own, as the anchored enzyme must remain functional in the new microenvironment of the cytosolic side of the compartment. To test if enzymes remain functional when attached to the compartment, another cargo construct was developed using the citramalate synthase enzyme from *Malus* x *domestica* (MdCMS). MdCMS catalyses the formation of (R)-citramalate from acetyl-CoA and pyruvate (Figure 5.8A) and natively localises to the chloroplast[176]. Citramalate was not expected to form in untransformed *N. benthamiana* leaves, thus allowing easy identification of enzyme activity using GC-MS to detect (R)-citramalate from the crude lysate of *N. benthamiana* leaves expressing MdCMS. An MdCMS-mC-SpC-his cargo construct was co-expressed successfully with both C22Y and C22Y-SpyT-his and localised as expected: in the former primarily to the chloroplasts (since the native localisation signal was not removed) and in the latter to the compartment (Figure 5.8B). A genetically fused variant (C22Y-MdCMS) was also successfully cloned but expression levels were too low to elicit compartment formation and thus were not used in later experiments.

GC-MS analysis of mature *N. benthamiana* leaves sequentially transformed with C22Y/C22Y-SpyT-his and MdCMS-mC-SpC-his clearly showed that both produced citramalate, while negligible amounts were detected in the blank and WT samples (Figure 5.9). This meant that the MdCMS enzyme remained functional even when attached to the surface of the compartment, opening the possibility for further experiments using the compartment as a platform for anchoring enzymes and use as a metabolic microdomain. Unexpectedly, the amount of citramalate produced was fivefold higher when co-expressed

with C22Y-SpyT-his and anchored to the surface to the compartment, than when co-expressed with C22Y and localised to the chloroplast (Figure 5.9). This was not due expected to be due to a difference in expression levels of MdCMS, as comparison of the genetically fused mCherry fluorescence showed no statistically significant difference ($p = 0.142$).

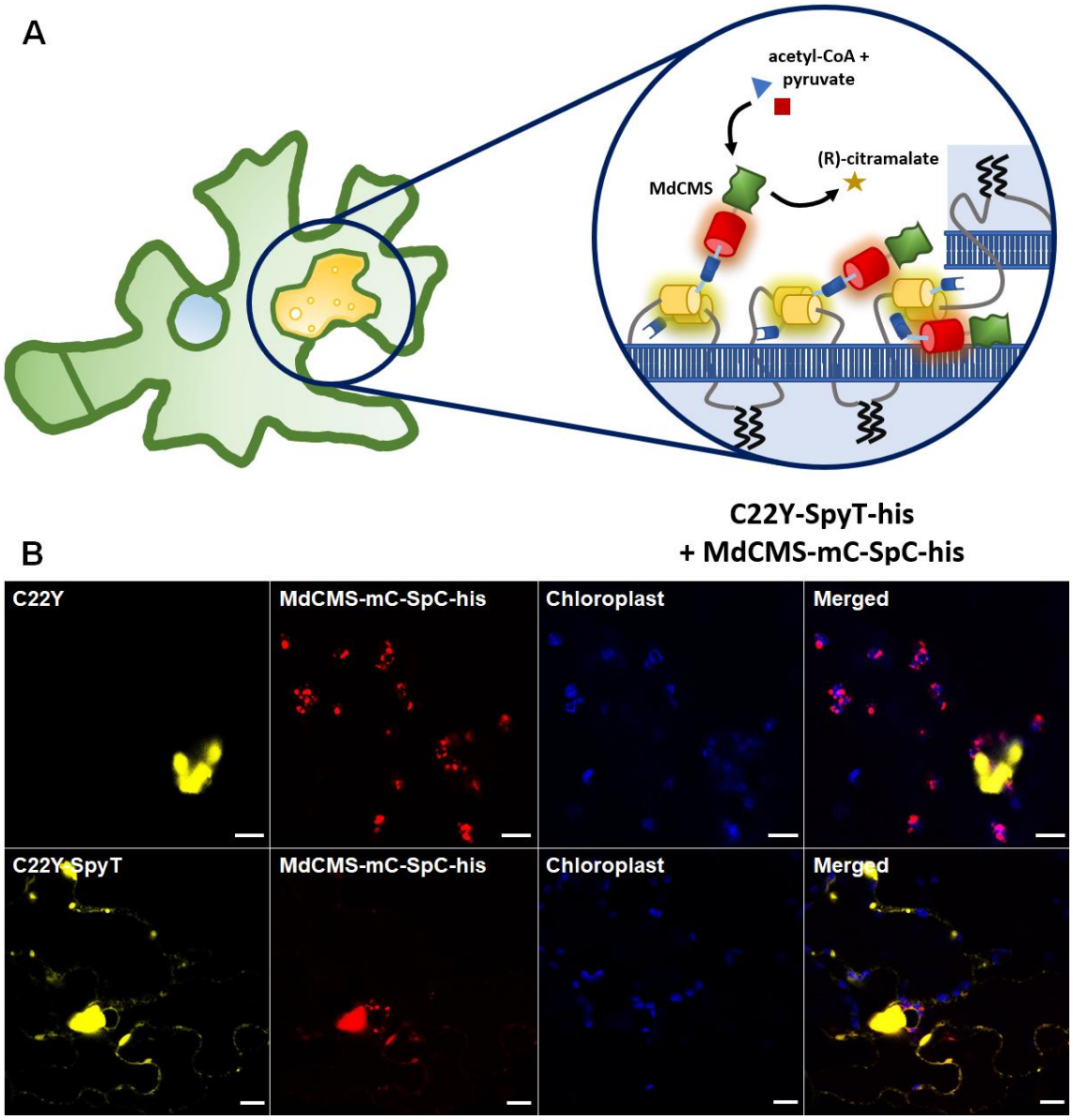


Figure 5.8 (above): Visualisation of anchoring an enzyme to the surface of the compartment. A: The schematic of the proposed experiment. MdCMS-mC-SpC-his cargo protein is covalently bound to the cytosolic side of the compartment via the SpyCatcher – SpyTag system, where it facilitates the production of (R)-citramalate from pyruvate and acetyl-CoA. The end product can be measured via GC-MS to test enzymatic activity. B: Fluorescent confocal microscopy images (yellow: YFP, red: mCherry, blue: chloroplast autofluorescence) of MdCMS-mC-SpC-his with C22Y/C22Y-SpyT-his in mature *N. benthamiana* leaves 7 d after *A. tumefaciens*-mediated sequential transient transformation. The C22Y/C22Y-SpyT-his constructs were introduced 14 d before the images were taken. Formation of smaller compartments were observed in rare cases during sequential transformation. Scale bars: 10 µm.

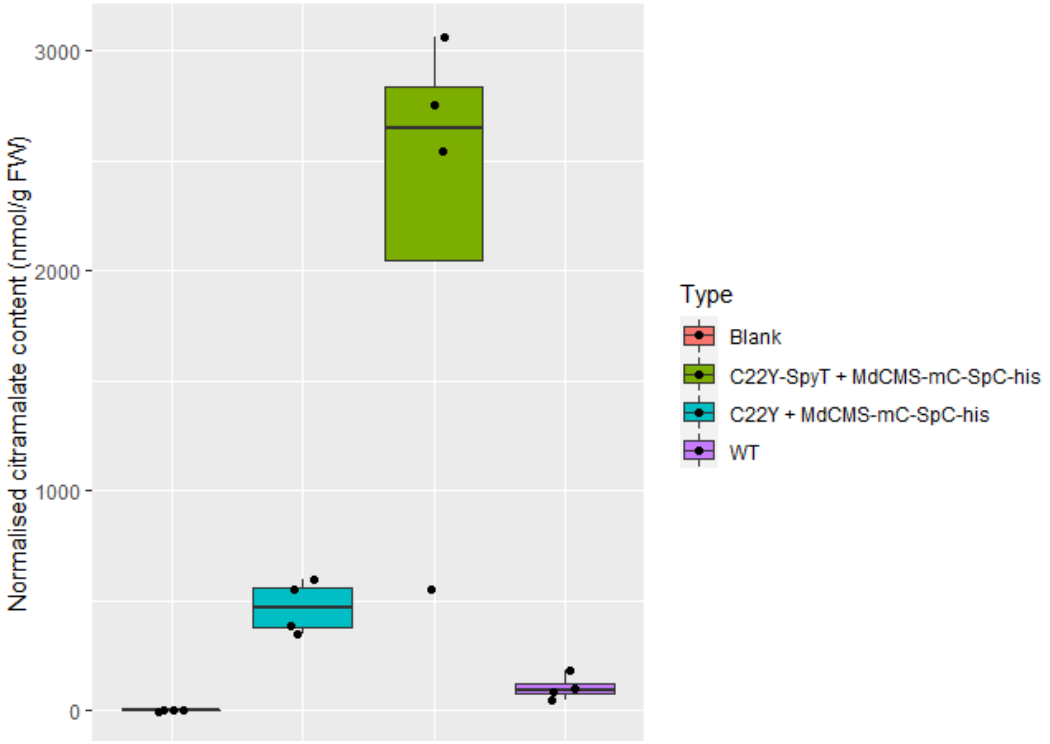


Figure 5.9: The activity of an enzyme anchored to the surface of the compartment. GC-MS analysis of samples from Figure 5.8, comparing the normalised citramalate content (nmol/g fresh weight) of *N. benthamiana* leaves 7 d after *A. tumefaciens*-mediated sequential transient transformation. The boxes represent the interquartile range, the horizontal line in the box shows the median, and the whiskers the minimum and maximum values (excluding outliers). n = 4 for all categories

5.2.6 Characterisation of stably transformed *A. thaliana* lines

The development of stably transformed lines which do not require constant *A. tumefaciens*-mediated transient transformation would be a great boon for practical application of the compartment system. The benefits are twofold: a stably transformed plant would express the compartments more uniformly in every cell, as the unpredictable plasmid insertion site into the genome of *A. tumefaciens*-mediated transformation causes protein expression level to vary between each transformed cell. Furthermore, stably transformed lines would considerably reduce the workload and simplify the production pipeline when working at scale. However, while previous work suggested that the compartments do not have a significant impact on plant health and disruption to the cells are minimal, only leaf epidermis was investigated. Stably transformed lines would have the genes for the compartment-forming constructs in all tissue types and under the drive of the strong constitutive cauliflower mosaic virus 35S promoter[202] are expected to express them throughout the whole plant, and thus the effects of this on the plants must be investigated before stably transformed lines can be used for any of the applications described above. While the use of a tissue-specific promoter was a possibility, changing of promoter from the 35S used in transient expression experiment would have meant a potential confounding factor for any observed differences, as it would have been unclear if the changed promoter (and thus possibly different expression levels) or the change in species, tissue type and developmental timing was the causal factor for any difference.

Stably transformed *A. thaliana* T1 C22Y and 22Y seedlings were selected based on antibiotic-resistance and inspected for fluorescence (Figure 5.10), which showed that every tissue type examined exhibited strong YFP signal. Four T2 lines were then isolated, which

also showed YFP fluorescence in all tissue types (Figure 5.10). Fluorescence was notably stronger in the roots of T2 plants than other tissue types.

T2 plants were further investigated using laser-scanning fluorescence confocal microscopy. Initial imaging of the leaf and stem showed ubiquitous presence of large compartments in both tissue types for both C22Y and 22Y expressing plants (Figure 5.11). Seeds were also imaged, showing YFP fluorescence with a different disposition in C22Y and 22Y seeds. 22Y-expressing seeds showed faint YFP fluorescence in the seed coat, but most of the signal was concentrated to the micropylar endosperm, with none present in the embryo. C22Y-expressing seeds had a more dispersed signal, with no distinct concentration of fluorescence in the micropylar endosperm, but some compartment-like structures present in the seed coat, and no signal in the embryo (Figure 5.11).

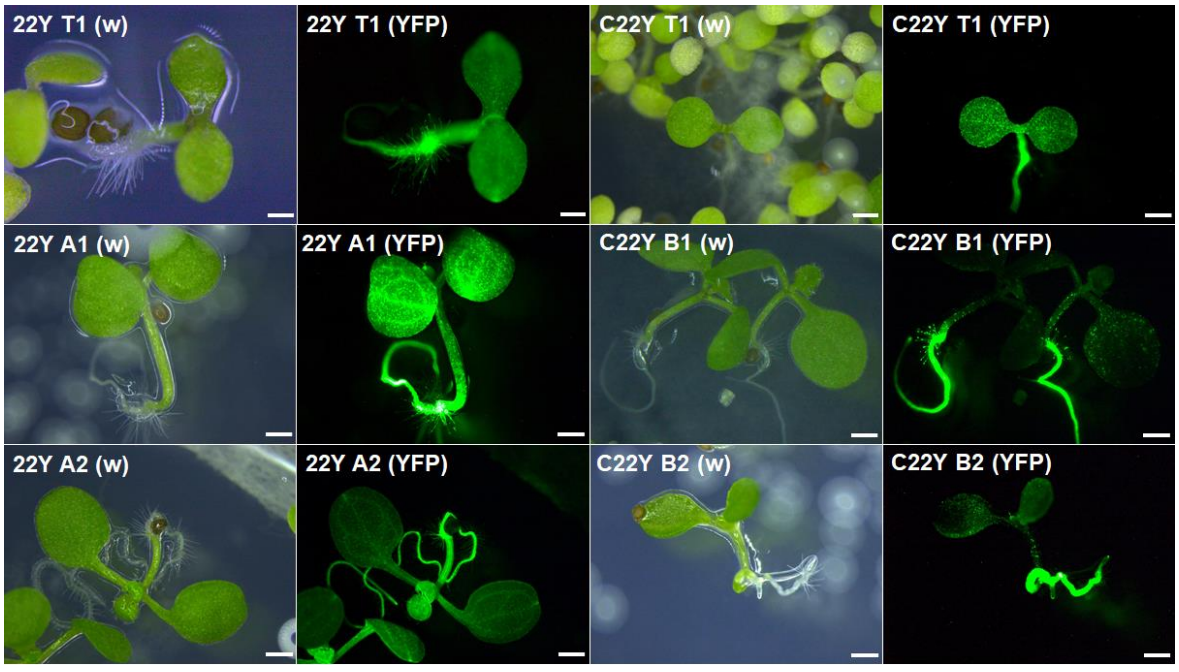


Figure 5.10: Fluorescent whole-seedling imaging of stably transformed A. thaliana. 5-6 d old seedlings on MS-media plates were imaged using white light (w) and for YFP fluorescence (YFP). The top row shows two pairs of images from T1 C22Y and 22Y lines, while the bottom two row show four isolated T2 lines. Fluorescent signal is present throughout all tissues, but notably, in most T2 lines the YFP fluorescence in the roots is stronger than elsewhere in the plants. Scale bars: 1 cm.

High-resolution airy-scanning confocal microscopy of T2 plant leaves and stem revealed the presence of the expected compartment types for both 22Y and C22Y, with the former showing elongated, strand-like structures and the latter large, globular compartments with a more even surface coverage of fluorescent signal (Figure 5.12), just as described in transient expression experiments in *N. tabacum* and *N. benthamiana*. There was no noticeable difference in the compartment structures in the stem when compared to the ones in leaf tissue.

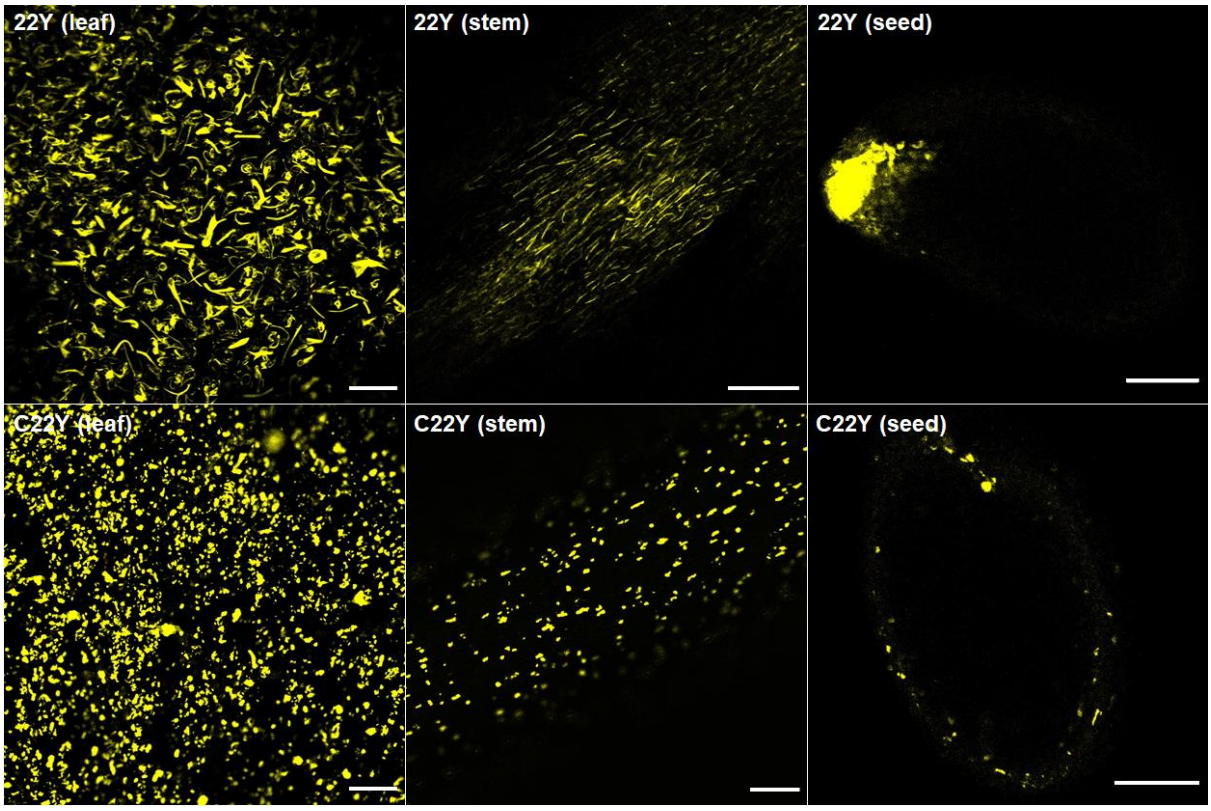


Figure 5.11: Confocal microscopy of T2 stably transformed A. thaliana lines. The leaf and stem of 6 weeks old plants were imaged for YFP fluorescence showing consistent presence of compartments in all cells for both 22Y and C22Y lines identical to the previously described compartment morphologies for these two constructs. Seeds from both lines were also imaged, showing a more specific distribution of YFP signal. Scale bars: 100 µm.

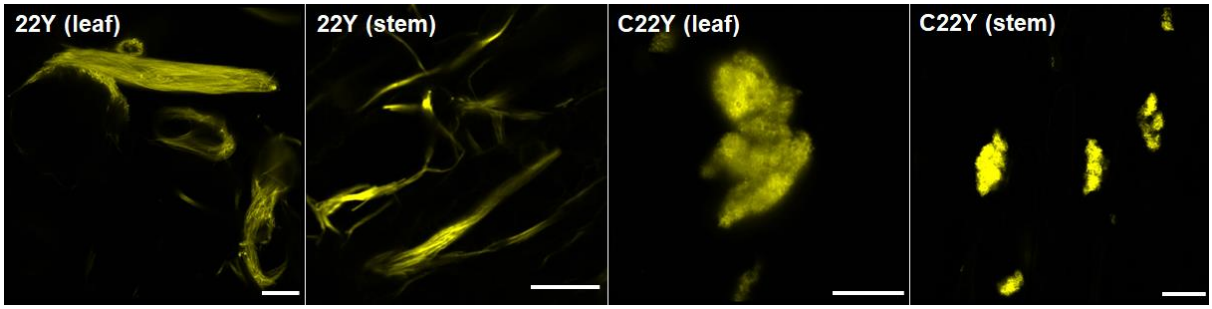


Figure 5.12: High-resolution airy-scanning confocal microscopy of T2 stably transformed A. thaliana lines. The leaf and stem of 6 weeks old plants were imaged for YFP fluorescence. Both compartment types show morphologies identical to the previously described transiently expressing compartments for these two constructs. Scale bars: 10 µm.

After confirming the presence of the expected compartments, the T2 lines were further characterised to assess if stable expression of the compartments caused any obvious macro-scale phenotypic changes. To this end, the leaf surface area of 4- and 5-week-old plants were measured by imaging them from above (Figure 5.13A) and analysing the images for green leaf area using the LeafLab software package (Mark Fricker, unpublished) (Figure 5.13B). Upon visual inspection, C22Y lines showed a clear pattern of accelerated growth when compared to WT or 22Y, with quicker growth of leaves and earlier start of flowering spike growth (first stalks appeared at week 3, while only at week 4 for WT). 22Y had a similar leaf growth rate to WT, but consistently slower flowering spike growth (Figure 5.13C). All lines set viable seed normally.

LeafLab was used to measure three key parameters of T2 *A. thaliana* at week 4 and week 5: leaf area, leaf convex area and solidity (Figure 5.13B). Leaf area is the most important of these three, since it is directly related to photosynthetic capacity of the plants. Note that by week 5, most plants started to have overlapping leaves, so by that point the measured value was a better estimate of effective photosynthetic leaf area, rather than total leaf

area. Leaf convex area and solidity gives a good insight into the shape and spread of the leaf coverage of the plants.

The results of the statistical analysis can be found in Table 5.1 and Figure 5.14. All four T2 lines had statistically significantly greater leaf surface area than WT at both week 4 and 5. Convex area was greater for 22Y A2 and C22Y B1 and B2 when compared to WT at both week 4 and 5. Solidity showed a different pattern, with the quicker growing C22Y lines having higher values at week 4, but by week 5 they have been overtaken by the other lines and 22Y A1 showing the highest values (and thus the most even coverage). This combined with the large convex surface area and large surface area suggest that C22Y lines were more likely to spread out with longer leaves than 22Y and WT plants (Table 5.1 & Figure 5.14).

Dry weights of the plants were measured at week 8. Total above-ground dry weight was less for 22Y A1 and C22Y B2 but not statistically different to WT for 22Y A2 and C22Y B1. Since most of the *Arabidopsis* mass at this point was from the stem, the same pattern was found stem dry weight measurements as well. 22Y lines had a greater leaf dry mass, and C22Y B1 lower than WT. This was in line with visual inspection, as 22Y lines had noticeably more green leaves at week 8 than the other lines, possibly due to the aforementioned observation of a developmental delay (Table 5.1 & Figure 5.14).

Since independent lines of each construct do not follow a clear and identical pattern for the measured phenotypes, more independent lines should be grown to assess the compartments' impact on stably transformed lines. However, these initial studies suggest that in line with transient expression studies, the compartments are not notably detrimental to the plants and generally normal phenotypes were observed.

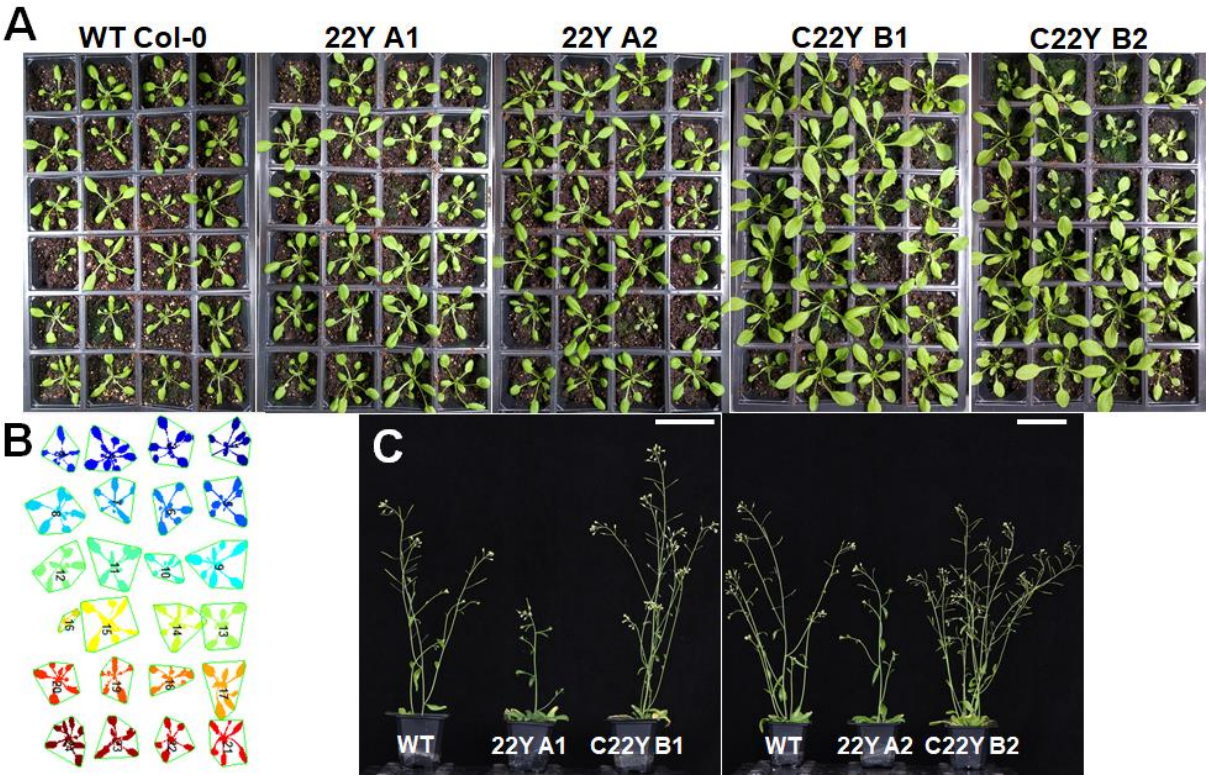


Figure 5.13: Stably transformed T2 A. thaliana phenotypes. A: Top-down imaging of T2 lines 3 weeks after transplantation to soil from MS selective media plates. C22Y lines show clearly quicker leaf growth. B: Example image (of WT Col-0 from A) of how the LeafLab image analyser software package differentiates between individual plants. The green outline shows convex area. Solidity is calculated as area (coloured) divided by convex area. C: Side-by-side comparison of 5 weeks old T2 plants. C22Y lines are larger than the WT or 22Y lines, possibly due to the earlier start of stem growth. 22Y showed delayed stalking compared to WT. Scale bar: 5 cm.

T-tests against WT (p-values)	22Y A1	22Y A2	C22Y B1	C22Y B2
Week 4 leaf area	0.0486 (*)	0.0011 (**)	7.4×10^{-8} (***)	4.6×10^{-5} (***)
Week 4 leaf convex area	0.1128	0.0015 (**)	9.2×10^{-7} (***)	0.0005 (***)
Week 4 leaf solidity	0.2459	0.9307	0.0130 (*)	0.0121 (*)
Week 5 leaf area	0.0304 (*)	0.0065 (**)	1.4×10^{-7} (***)	0.0008 (***)
Week 5 leaf convex area	0.1146	0.0055 (**)	1.1×10^{-8} (***)	5.3×10^{-6} (***)
Week 5 leaf solidity	0.0314 (*)	0.6458	6.1×10^{-6} (***)	0.0068 (**)
Total above-ground dry weight	0.0014 (**)	0.3630	0.2581	0.0240 (*)
Total leaf dry weight	0.0059 (**)	3.0×10^{-5} (***)	0.0002 (***)	0.4388
Total stem dry weight	0.0001 (***)	0.6512	0.5170	0.0204 (*)

Table 5.1: Results of statistical analyses of key macro-scale phenotypes. LeafLab was used to determine leaf area, leaf convex area and leaf solidity of T2 lines at week 4 and 5. The dry weight of leaves and stems were measured at week 8. Asterisks show significance (* for $p < 0.05$; ** for $p < 0.01$; *** for $p < 0.001$). n = 24 for week 4 and 5 measurements, n = 20 for dry weight measurements for all categories. Due to time-constraints, this experiment was only carried out once, and should be repeated to confirm the results.

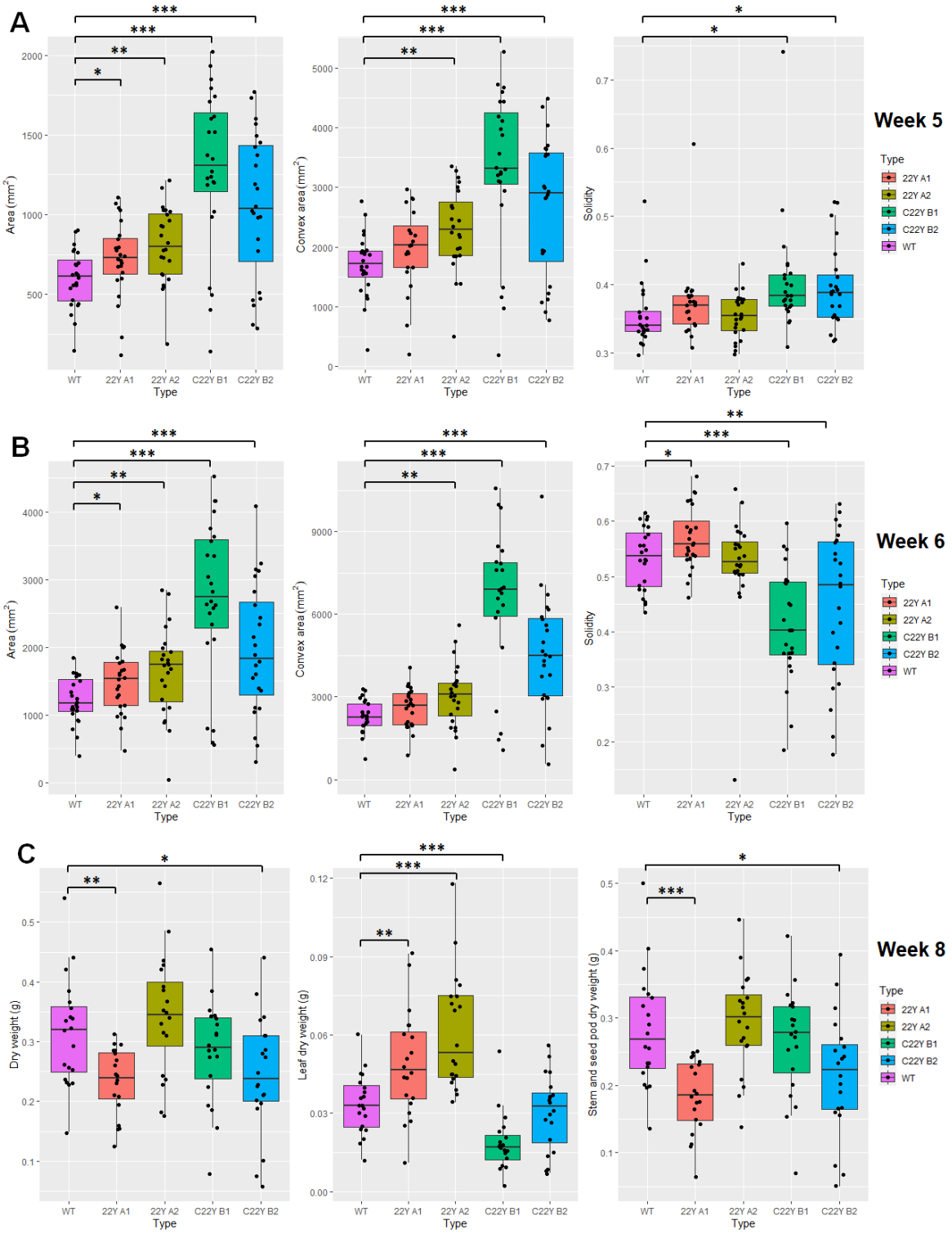


Figure 5.14: Graphical representation of the statistical analysis of the key macro-phenotypes. A: Week 4 leaf surface area analysis via LeafLab. n = 24 B: Week 5 leaf surface area analysis via LeafLab. n = 24 C: Dry weight measurements of leaves, stem and total above-ground biomass taken at week 8. The boxes represent the interquartile range, the horizontal line in the box shows the median, and the whiskers the minimum and maximum values (excluding outliers). n = 20 for all categories

5.3 Discussion

5.3.1 The viability of a synthetic ER-derived compartment as a platform for improving plant-based recombinant protein production and as a metabolic microdomain

Several of proof-of-principle studies were carried out successfully to begin the first steps towards the application of the compartment system as a platform for both the production and purification of recombinant proteins and as a metabolic microdomain. Protein purification was tested using two different methods (his-tag based affinity purification and immunoprecipitation) on a number of targets, including compartments with genetically fused and covalently attached enzymes. Enzymatic activity on the cytosolic surface of the compartment was also confirmed. Finally, the establishment and initial characterisation of stably transformed *A. thaliana* lines with constitutively expressed compartments in all tissues opens up the possibility of moving beyond *A. tumefaciens*-mediated transient expression as the only method to induce the formation of compartments.

When building a protein-induced compartment, the proportion of the scaffold molecules inside the compartment is a key metric of specificity. This is especially important when functionalisation depends on attachment of (cargo) proteins to the surface of the compartment via direct interaction with the compartment-forming scaffold (either via genetic fusion or covalent bonds). If a high proportion of the scaffold molecules are outside the compartment, then functionalisation will be hampered as the proteins of interest (cargo) will be less concentrated to the compartment.

OSER studies in general do not to provide a measurement of this metric, since most of the OSER research focuses on the formation and characterisation of OSER structures, rather

than functionalisation[117,140,143,148], so direct comparison is not possible. As such, this might be the first reported measurement of proportion of scaffold molecules inside an OSER. Comparison to other types of compartments is possible however, where data is available. For example, a recent study on a phase-separated synthetic membraneless organelle showed that depending on the type of scaffold protein used, between 95% and 72% of them were inside the compartment[66]. This is close to the values reported with my membranous compartment system (92-82% - Figure 5.1B). The same study also reported similar cargo capture rates, with values ranging between 91% and 50%, depending on the combination of host chassis, scaffold and cargo molecules[66]. Other phase-separation studies have shown comparable cargo-capture rates (e.g., 95-80% in one study[63]), and our membranous compartment did too (87% and 49% - Figure 5.5).

This difference between the capture rates of the two cargo types described in Figure 5.5 is likely due to the different transformation methods used. HbpA (with 87% capture rate) was expressed after the compartments have already formed using sequential transformation, while Luc (49% capture rate) was co-expressed with C22Y-SpyT-his simultaneously. The rationale is that if the compartment has already formed, the cargo has a higher chance of encountering and covalently binding to scaffold proteins due to their high abundance, reducing the risk of transport to an inaccessible location (e.g., the vacuole) and allowing capture of the cargo before expression might be downregulated. Since the compartment was shown to be stable for up to 6 weeks (Figure 3.5) this allows sequential transient transformation to be a viable strategy, by introducing the compartments early, and re-transforming the same leaves later with the desired cargo construct. This strategy expands the flexibility to use this platform by reducing time constraints and thus increases ease-of-use when working with mature plants.

Flexibility is also increased by the introduction of the SpyCatcher – SpyTag system to allow a covalent binding of the target cargo proteins. With this system, only a single, well-characterised compartment-forming scaffold needs to be optimised (e.g., C22Y-SpyT-his) and cargo types can be interchanged as required. This is generally preferable over genetic fusion of target proteins to the N- or C-terminus of the compartment-forming scaffolds, since the effect of the fusion protein on compartment formation and structure must be assessed for every construct. It also reduces the number of molecular cloning steps, decreasing the time needed to go from an experimental plan to an expressed protein. Genetic fusion can also have a major effect on the expression levels of a new construct: for example, when attempting to test the enzymatic activity of MdCMS on the cytosolic surface of the compartment, a C22Y-MdCMS genetic fusion construct was cloned, but was never successfully expressed to a high enough level to induce compartment formation. However, the cargo capture method worked successfully (Figure 5.8).

An alternative to sequential transformation is to use stably transformed C22Y-SpyT-his plants that express compartments, then introduce (via *A. tumefaciens*-mediated transient transformation) the desired cargo proteins. The initial characterisation of stably transformed *A. thaliana* lines showed that compartments can be formed and are stable. Thus, generation of C22Y-SpyT-his plants should now be the recommended next step to add the flexibility of post-translational covalent binding of cargo proteins and enable purification of them. For practical applications, this combined strategy of stably transformed compartment-expressing plants, with transient introduction of the desired cargo might be preferable over the sequential expression strategy, as it is even more streamlined, and reduces the potential risk of complications from plant immune reactions to repeated *A. tumefaciens* infections[203]. Purification yield of the constructs was between

22 – 56% when using his-tag purification, and further optimisation of this method should focus on improved solubilisation of the membrane proteins and completely eliminating the dimerization capacity of the YFP during purification. Possible strategies could include the modulation of pH, salt concentration and detergent strength in the lysis buffer.

Attaching an enzyme to the cytosolic surface of the compartment also proved successful. Indeed, MdCMS produced more citramalate when bound to the compartment, than when localised to the chloroplasts (Figure 5.9), opening the path for further similar experiments in order to eventually produce a multi-enzyme assembly for probabilistic metabolic channelling. However, the key takeaway should be functionality, not improved enzyme efficiency as this was only a single enzyme proof-of-principle study on a limited number of samples and the cause of difference in citramalate production is unclear. A recent discovery of evidence for metabolically separate pools of pyruvate (one of the substrates of MdCMS) in plant cells[204] hints at an alternative explanation, that substrate availability could also be a causal factor, and the different subcellular microenvironment of the cytosol compared to the chloroplast could also play a role. It is important to note that the citramalate levels produced here are below what a previous study found when transiently expressing MdCMS in *N. benthamiana*, where up to 30 µmol of citramalate was produced per g fresh weight[176].

In general, the proof-of-principle experiments described above showed great promise for the functionalisation of the synthetic ER-derived membranous compartment. A wide variety of functions are enabled through the development of a reliable purification method and the *in vivo* covalent binding of cargo proteins and functional enzymes to the surface of

the compartment. Future goals and potential further improvements to this platform are discussed in Chapter 6.

5.3.2 The compartment's potential protective capacities for attached proteins from proteases

Throughout the previous chapters, hints have accumulated that suggest the possibility that compartmentalisation into our synthetic ER-derived membranous compartment might protect proteins from cleavage by certain proteases. Compartmentalisation isolates the proteins from the bulk of the cell, reducing the rate of exchange between the compartment and both the rest of the ER lumen and the bulk cytosol (Figure 3.9), which could offer some protection from proteases. The compartments also have a very long lifespan, as even 6 weeks after transient transformation it was observed (Figure 3.5), so the turnover rate of the scaffold proteins is expected to be low. Stably transformed lines also express the compartments throughout the life-cycle of the plants (Figure 5.11).

The hypothesised ER – peroxisome – vacuole pathway described in Chapter 4.3.2 might also support this hypothesis. Without the formation of a compartment, the defective (non-compartment forming) scaffolds are removed through the peroxisome into the vacuole. One explanation was that this could be a mechanism for the removal of excess or misfolded proteins. Since compartment forming scaffolds do not follow this pathway, proteins in the compartment might be protected from this removal and degradation.

Finally, Figure 5.6 offers another piece supporting evidence. There are lower molecular weight bands present in the Western Blots that were identified as HbpA/C-mC-SpC-his proteins with some N-terminal parts of the HbpA/C cleaved off. These bands did not undergo the band-shift expected from the C22Y-SpyT-his binding, despite retaining their

SpyCatcher. It is possible that once the cargo is bound to C22Y-SpyT-his, this N-terminal cleavage is no longer possible, hence no smaller band-shifted signal is present that would correspond to the cleaved cargo interacting with C22Y-SpyT. Since the SpyTag – SpyCatcher binding goes almost to completion for sequential expression (Figure 5.5), any surplus unbound cytosolic cargo could be cleaved by the proteases.

Our current knowledge of plant ER-associated degradation (ERAD) mechanisms could also lend support to this theory, since many proteins involved in ERAD are ER membrane proteins, which could be diluted in the compartment due to the large quantity of scaffold proteins present. Furthermore, ERAD requires retrotranslocation of proteins targeted for degradation, which could be blocked by the complex oligomerisation of the scaffold proteins[205,206]. As such, the compartment system might interfere with both of these mechanisms and thus ERAD.

Taken together, these clues suggest the possibility of a potential protective capacity of the compartment against proteases: however, it will require further testing to assess this hypothesis. This could utilise anchored cargo proteins with degradation tags: N-terminal aromatic and basic amino acids are common targets of the *A. thaliana* PRT1 and PRT6 E3 ligases which initiate degradation of recognised proteins via the proteasome[207–209]. Plants expressing these constructs can be treated with a protein synthesis inhibitor (e.g., cycloheximide[210] or cylindrospermopsin[211]) and the change in the fluorescent signal of cargo proteins can be observed to determine the half-life of the proteins. If indeed, the compartment offers protection against degradation, then anchored cargo proteins are expected to have a longer half-life than unbound cytosolic variants. An alternative test could utilise luminal Protein A fusion to the compartment and using antibodies as cargos

to determine if recovery of these valuable proteins can be increased by sequestering them from the bulk of the cell[183].

5.3.3 Phenotypical differences of stably transformed *A. thaliana* lines and their consequences for application

The stably transformed C22Y and 22Y expressing *A. thaliana* T2 lines were all shown to have a greater leaf area than WT (Figure 5.14A-B). Developmentally they were less uniform, with C22Y growing slightly longer leaves as seen by the higher convex area and lower solidity by week 5. While the greater leaf area and consequentially greater photosynthetic capacity could have explained the increased growth of C22Y, it is in contrast with the delayed stem growth of 22Y lines which also had greater leaf area (Figure 5.14). For the dry weight of plants, no clear pattern emerged, with some lines having lower total above ground biomass than WT, while others did not show a statistical difference (Figure 5.14C).

Overall, the introduction of the stably expressing compartments had some effect on *A. thaliana* phenotype, but the consequences for practical application are not clear yet. The quicker growth of C22Y is beneficial, although the greater spread of leaves will require slightly more space for the plants. The increased leaf area for C22Y and 22Y could be beneficial, since most of the protein production or metabolic engineering (via transient transformation of cargo proteins) would take place there. However, the most important conclusion is that the massive enlargement of the ER in the compartment-expressing lines does not appear to have any major detrimental effect on plant growth or reproduction in the stably expressing lines, in line with earlier findings on plant health for transiently

transformed lines. Thus, functionalisation of stably transformed lines appears to be feasible.

Investigation of the roots of stably transformed lines might shed more light into the source and consequences of these phenotypical differences. Since root tissue was never transiently transformed to generate compartments, we do not have any indication of the effect it might have, although fluorescent whole seedling imaging showed a strong signal in the roots of T2 stably transformed *A. thaliana* (Figure 5.10), and OSER compartments have been described before in the roots of *A. thaliana*[208], so analysis on the impact on the cells of the root (e.g., by AnalyzER) should be possible.

Chapter 6: Discussion

6.1 Short summary of results and future perspectives on the compartment system

During the course of this project, starting out from a single synthetic protein that induced the formation of a structure similar to OSER, I have developed a range of compartment-forming constructs that give rise to several different ER-derived compartment morphologies. After characterising these, I have presented a hypothesis about the driving forces for the formation of these ER-derived synthetic compartments. According to this hypothesis, compartment formation requires the expression of a protein at sufficiently high concentration, targeted to the ER-membrane with a cytosol-facing antiparallel oligomerising domain. The exact structure of the oligomerising domains and other peptide tags is hypothesised to determine the final compartment structure and form. As broad similarities exist in this model with previous explanations for OSER formation, the compartments I have described in this thesis are likely a subtype of OSER. However, a lack of a universally agreed definition of OSER makes it difficult to precisely categorise the compartments.

The properties of these compartments were investigated to establish their suitability for functionalisation. Notably, the membranous compartment had a diffusional barrier between the ER and the cytosol but had little impact on the rest of the ER of the host cell and also on the whole plant in terms of growth and morphology. I demonstrated that a single, easily modifiable scaffold protein is sufficient to induce the formation of the compartment. These properties made the compartment attractive for functionalisation

with applications such as purification of expressed proteins and as a tool for metabolic engineering, and several proof-of-principle experiments were successfully carried out.

In this chapter I discuss what further investigations and improvements would be beneficial to further understand and exploit the compartment system. In particular, I discuss remaining open questions and potential avenues to answer these regarding the basic properties of the compartments, including in-depth *in silico* modelling of the constitutive scaffolds and the benefits of expanding the potential range of constructs, investigative techniques, and host organisms. I also discuss the potential commercial applications, such as production of difficult-to-express proteins, exploring the possibility of building a platform for synthetic metabolic microdomain, and the long-term future improvements to the system that would enable further applications. Finally, I compare this compartment system to other available compartmentalisation strategies, to determine if it has a viable niche in the plant synthetic biology field.

6.2 Potential further investigations into the basic properties of the ER-derived synthetic compartment system

6.2.1 Incorporating the advancements in *in silico* modelling of proteins and protein-lipid interactions into the design pipeline of compartment-forming constructs

A great deal was uncovered about the structure and formation of the compartment system over the course of this project. However, there are still a number of outstanding questions. Knowledge gained from, and *in silico* methods developed for the investigation of other membranous structures could be applied to leverage more information on the structure and formation of the compartment system in the future.

As the number of compartment-forming constructs I investigated increased, four distinct compartment morphologies were described. A hypothesis was presented which described the relationship between scaffold composition and the ultrastructure of resulting compartments based on the properties of the oligomerising domains (Figure 4.8). While this explains the results of this project, it might prove difficult to extrapolate this framework to new constructs that would be designed in the future to predict their compartment ultrastructure.

Advances in *in silico* modelling of protein-lipid interactions in the ER membrane could improve our understanding of how scaffold structure affects the compartment ultrastructure. For example, Shemesh *et al.* (2014) developed a comprehensive model of how a small number of ER-shaping membrane proteins can cause the formation of all known ER morphologies. They used the local concentration of curve-inducing ER membrane proteins to model expected ER shapes, and many of their findings might be

applicable to my constructs and compartments[209]. According to their model, a higher concentration of curvature-inducing membrane proteins promotes the formation of tubules which might correspond to the tubular 22Y structures. Conversely, a lower proportion of curvature-inducing membrane proteins allow more sheets and even helicoid stacked structures to form, potentially resembling the C22Y and G22Y structures, respectively[209]. Measuring local concentrations of these scaffolds could unveil if there is a connection, as a higher local concentration of 22Y would be expected when compared to C22Y and G22Y. If the model is validated for my ER structures as well, it could be used to aid the design of future compartments. However, the Shemesh model did not incorporate the effects of inter-membrane oligomerising domains might have on the structure of the ER membrane and introducing these effects into the model might improve the predictive power of the model for the compartment structures described in this thesis.

Recent advances in the field of ER-shaping membrane proteins could also be relevant to the study of the compartment system described in this thesis. For example, reticulons are oligomerising transmembrane ER proteins that induce tubule formation by wedging (insertion of multiple long transmembrane domains to displace lipids and cause a kink in the membrane bilayer) and scaffolding (within-membrane oligomerisation)[108,110,112,210,211]. Both of these forces could be in play during the formation of the compartment as well, owing to the compartment-inducing scaffold constructs having a broadly similar set of domains: transmembrane domains for wedging and oligomerising domains for scaffolding. This could explain why these constructs have such a great effect on ER rearrangement, despite the fact, that none of the individual constitutive domains are involved in ER structure formation. Another two sets of ER-shaping transmembrane proteins, atlastins, such as the plant orthologue RHD3[114,115], and the Lunapark proteins[211] also share

similarities with my constructs, with oligomerising domains cross-linking membranes while being anchored by a transmembrane domain[212]. However, overexpression of these ER-shaping proteins does not induce the formation of OSER-like structures: overexpressing reticulons leads to the constriction of the ER lumen[112,213], Lunapark overexpression broadly expands the cisternae[114,213], and RHD3 overexpression does not cause ER-rearrangement in plant cells[114,214]. This suggests that even if the compartment-inducing scaffold constructs are structurally similar to these ER-shaping proteins, their mode of action on ER restructuring is likely different. With the recent breakthroughs in atomic accuracy modelling of proteins[218,219] and the increasing interest in modelling lipid-protein interactions, it might soon be possible to directly test the similarities and differences between my scaffold proteins and the atlastin, reticulon and Lunapark families of ER-shaping proteins to inform future scaffold design.

Overall, incorporating *in silico* modelling of proteins and protein-lipid interactions into the pipeline of compartment-forming construct design might refine the existing model of the relationship between scaffold proteins and compartment ultrastructure and formation. Furthermore, this could allow us to predict the compartment phenotypes of new scaffold constructs prior to cloning and recombinant protein expression.

6.2.2 The use of novel construct domains and techniques to investigate compartment formation and structure

Beyond modelling, more can be learned about compartment formation and structure by expanding the portfolio of available compartment-forming constructs, by introducing new types of domains, and by using further techniques to characterise the compartments.

A number of different oligomerising domains were tested during this project, but coverage of this topic was by no means comprehensive. For example, while tetrameric domains were found to be suitable for compartment formation, many other common higher order oligomerising domains are yet to be tested, such as trimers or hexamers (e.g., the synthetic coiled coil domains CC-Tri[216] and CC-Hex[217]). Hetero-oligomers are also a potential avenue of future research. These domains would increase the complexity of experiments since multiple constructs would be required for the formation of a compartment, preferably at similar expression levels. However, such constructs could be of interest for the formation of multi-enzyme assemblies, with a more defined intra-compartmental spatial organisation. In theory, the different constructs of a hetero-dimer could each have a different enzyme of the same metabolic pathway attached, ensuring that a pair of these enzymes would always be in close proximity within the compartment. However, introduction of a pair of large enzymes could sterically interfere with compartment formation, so the use of hetero-oligomers for this purpose must be tested carefully first.

One domain which was never replaced in any of the compartment-forming constructs during this project was the transmembrane domain BP22. As such, the effect of the transmembrane domain on the formation and structure of the compartment was never investigated beyond the removal of BP22 from the cytosolic G-Y(cyt) and mC-FDH-his constructs, which did not form a compartment. Transmembrane domains have been long known to affect proteins in many ways: they affect subcellular localisation[218], can be involved in retention and degradation in the ER[219] and are directly involved in shaping the membrane (e.g., reticulons)[112,210]. Many commercially valuable proteins, or enzymes of interest that could be anchored to the surface of compartments also have transmembrane domains. As such, a deeper understanding of the relationship between constructs with

multiple transmembrane domains and compartment structure will be important for functionalisation, as well as the further characterisation of compartment formation and structure.

Other than testing new constructs, it is possible to expand the use of some of the methods used in this project, as well as introducing novel ones to leverage more information from the existing compartment-forming scaffolds. One such option is the further use of AnalyzER[219] to determine the effects of continuous expression of compartments on stably transformed plants, and other tissue types beyond mature leaf epidermis (such as roots and stem). This could provide clues on the underlying factors for the phenotypic differences described in Chapter 5 between WT and stably transformed compartment-expressing *A. thaliana.* Further improvements to AnalyzER might enable quantitative analysis of the compartments themselves to further characterise the morphology and kinetics of these structures.

Electron microscopy is another tool that was used in Chapter 3 to elucidate the membrane structures of 22Y and C22Y compartments, and its use should be expanded to other constructs, especially with the finding that the densely packed membrane structures of C22Y are likely below the resolution limit of airy-scanning confocal microscopy. As such a much more in-depth understanding of the internal structure of the compartments should be acquired using TEM, focusing especially on C22Y-SpyT-his, the best candidate for functionalisation.

A more comprehensive study of the compartments using TEM could provide further clues on the relationship between OSER, and the compartments described here, and possibly confirm the hypothesis that the compartments are a specific subgroup of OSER. TEM

images can also be used for computational analysis using SPIRE, a recently developed TEM image analysis tool, which was first tested on the highly organised membranes of etioplasts of several angiosperms[220]. SPIRE can help the identification and mathematical definition of the structural and spatial properties of complex membrane systems, and thus its output might be fed back into the previously proposed *in silico* modelling pipeline for an improved prediction of compartment morphology based on scaffold construct.

In future studies, other methods developed for characterising membranous systems could be applied as well. For example, membrane interactions within the compartments might be investigated by techniques that were used for the investigation of membrane contact sites (MCS). In essence, the compartment-forming constructs that connect separate ER membranes together can be viewed as a synthetic ER-ER homotypic MCS, since they tether membranes at close proximity without their fusion[221]. A recent review by Scorrano *et al.* (2019) summarises the various experimental approaches to study MCS, with many of these having direct relevance to the compartments as well[222]. For example, electron tomography for a complete 3D reconstruction of the reorganised ER membrane structures[223] or Förster resonance energy transfer (FRET) based or split-fluorescent reporters to detect the proximity of two membranes[222,224].

6.2.3 The possibility of exploring other hosts for the compartment system

OSER was first described in detail in mammalian cell lines[31,118], and since then, most of the work in this field used yeast or animal cells as hosts[143], and while plant OSERs are similar, there are some phenotypical differences[143]. Since the compartment system developed in this project only used plants as hosts, it would be a logical next step to explore yeast or mammalian cell lines expressing similar compartment-forming scaffold constructs. This

would allow the comparison of compartment structures from similar constructs between different host systems and help determine how generally applicable the findings of this project are beyond plants. With more data on compartment morphology from different kingdoms, the relationship between OSER and the compartment system can be further clarified as well.

Furthermore, developing the compartment system in new hosts can expand the range of potential applications. For example, expressing therapeutic proteins in plants is hampered by their orthogonal N-glycosylation. While this issue can be solved by genetic engineering[225], this can be difficult, and could be avoided by using mammalian cell lines for protein production.

6.3 Exploring the established and potential further applications of the compartment system

6.3.1 Production of difficult-to-express proteins

The compartment's large surface area and the high local concentration of scaffold proteins that are sequestered from the bulk of the cytosol raised the possibility of its use as a platform for the production of difficult-to-express proteins. The large membranous surface is especially important since it was suggested that one of the main bottlenecks to recombinant membrane protein production is the available membrane for accommodating these[226]. This idea is supported by a study with yeast strains with massive ER proliferation triggered by derepressed ER membrane biosynthesis. The researchers found that the yields of eight membrane proteins were increased when compared to expression in wild type yeast[227].

Membrane proteins and other proteins which generally have a lower expression level could be attractive targets, especially with the putative protective capacity of the compartment from proteases that was raised in Chapter 5.3.2, potentially further improving yield. Improving the expression of toxic proteins might also be possible due to them being compartmentalised and shielded from the bulk of the cell, reducing their harmful effects. This sequestration of toxins was previously suggested as one of the likely native functions of OSER in plants[194]. Toxic proteins could be genetically fused to the compartment-forming scaffolds or be attached to fully formed compartments (via sequential expression or stably expressing lines) as a cargo. For cargo-based attachment, post-purification cleavage from the compartments is required to recover the individual target proteins. This can be

achieved via the Tobacco Etch Virus (TEV) protease, since a TEV protease cleavage site is built-in to the SpyTag – SpyCatcher used in this project[228].

Expression of non-native proteins also often carry the risk of heterologous subcellular localisation, which can lead to degradation of mis-localised proteins. For example, N-glycan composition and analysis of final localisation showed that KDEL tagged proteins meant for ER retention often localise unpredictably[43]. Since the scaffolds were shown to almost exclusively localise to the compartment (Figure 5.1), this issue can be reduced when using this platform for protein production.

Exploring other purification methods of the compartment-forming scaffold proteins and of the commercially valuable cargo proteins is another area where improvement of this system is possible. A range of different purification methods could be assessed for their suitability. More established methods of interest include microsome preparation which was shown to be suitable for recovering plant ER fragments[229], sucrose density gradient-based methods which employ differential centrifugation to fractionate cellular content[230], or free-flow electrophoresis[234]. More recent advances in the field of membrane protein purification should be continuously assessed as well, such as the novel styrene-maleic acid method which focuses on the recovery of membrane proteins in lipid particles[232], and the simple "teabag" purification method which can in principle reduce purification time significantly[233].

Beyond the proof-of-principle protein purification and cargo attachment experiments described in Chapter 5, preliminary work was started on the expression of the monomeric toxic OmpG porin from *E. coli*[234], to test whether the compartment afforded increased capacity for toxic membrane protein production. Unfortunately, time-constraints imposed

by the COVID-19 pandemic meant that this work could not be finished, although constructs with OmpG genetically fused to C22Y and a SpyTag-fused OmpG variant was cloned. A large number of other commercially interesting targets were considered for these experiments, all of which could be used in future research into this area. Examples include the α-haemolysin protein from *Staphylococcus aureus* used by Oxford Nanopore Technologies for nanopore sequencing[235], ER-membrane bound cytochrome p450 enzymes which are of interest for the pharmaceutical industry and biocatalysis, currently mostly produced in *E. coli* and yeast often with great challenges[236], and the human Epidermal Growth Factor Receptor, a difficult-to-express membrane protein valuable for medical research[237].

6.3.2 Building a platform for a synthetic metabolic microdomain

Another potential application that was considered in my work was to use the compartment as a synthetic metabolic microdomain with the subsequent enzymes of a metabolic pathway concentrated to the surface of the compartment to improve metabolic flux towards a desired product. The diffusional barrier to the bulk cytosol (Figure 3.9 & Figure 4.5) also meant that intermediate metabolites of the target pathway are less likely to diffuse out of the compartment, which is beneficial when there are competeing pathways or the intermediate is toxic or unstable. Finally, since the size of the microdomain also directly influences probabilistic metabolic channelling[99,238], the large size of the compartment is also beneficial. Castellana *et al.* (2014) calculated that the efficiency of a metabolic pathway in a microdomain would peak between approximately 5 – 25 µm, close to the size of the compartments described in this thesis, although it is important to note that they used a more spatially segmented microdomain for their model, so the applicability of these calculations ought to be verified[98].

A simple way to determine if probabilistic channelling can be achieved in the compartment is to use a branching metabolic pathway where the intermediate metabolite is used by two competeing enzymes[83]. The 2,2'-biphenol breakdown pathway was selected to test the compartment's capacity to serve as a microdomain (Figure 5.7), but as no products or intermediates were detected using GC-MS, it is likely that the enzymes were not functional. The cause of this is unclear, although the cleaved protein constructs discovered in Figure 5.6 suggest that the enzymes HbpA and HbpC might be susceptible to some plant proteases. However, this pathway still presents one of the best possibilities for assessing the compartment's potential as a synthetic microdomain and as such could be pursued as a future experiment.

Nonetheless, there are other metabolic pathways as well that could be of interest for the testing of the compartment as a metabolic microdomain. The phenolic-rich tannins are a group of flavonoids primarily serving as defence molecules. Their biosynthesis has toxic intermediates and their production in honeysuckle petals was suggested to take place in dedicated OSER structures to protect the bulk of the cell from these[194]. As such, reconstituting this pathway on the surface of the compartment system could be an alternative way of testing the capacity of the compartment for metabolic channelling.

Another application of this system as a microdomain could be to use as a methanol detoxification platform. In plants, there is a methane capture system currently under development (Verena Kriechbaumer, Oxford Brookes University, Oxford, UK, personal communication) which allows plants to sequester methane from the ground and convert it to methanol. However, continuous low-level exposure to methanol is toxic. Methanol could be turned into CO_2 by the three enzymes methanol dehydrogenase, aldehyde

dehydrogenase and formate dehydrogenase. If these are compartmentalised, the bulk of the cell would be shielded from toxic intermediates (e.g., formaldehyde) and the released carbon dioxide could potentially also improve Rubisco efficiency by suppressing photorespiration[242], a major target for improving plant production.

There are many other metabolic pathways with potentially commercially valuable metabolites and branching pathways that would be suitable for this purpose, including biosynthesis of the grapefruit aroma terpenoid (+)-nootkatone[243], various pigments with important medicinal properties from the betalain family[244,245], and some flavonoids such as the antimicrobial kaempferol, which is of interest for its medical uses[246]. Importantly, recent experiments in yeast increasing the ER surface area noted an improved rate of terpene production[247,248], which raises the possibility that ER surface area expansion (which is the case for this ER-derived compartment system) could be used in plants as well to improve bioproduction of valuable compounds[249].

When preparing for experiments assessing the compartment for metabolic channelling it is important to keep in mind that when expressing multiple enzyme constructs, a significant difference in expression levels of the enzymes is possible and could have a marked confounding effect on the channelling experiment. This could be avoided to a substantial degree by the use of the recently developed IntF2A system for *N. benthamiana* polyprotein constructs[250]. This system uses a self-excising domain composed of a high-efficiency self-cleaving intein and a foot-and-mouth disease virus 2A peptide (F2A) to express polyproteins which later separate into the constitutive proteins which are then present at comparable concentrations levels[250]. As such, it would be more certain that differences seen in fluxes between competeing pathways is not just due to differences in enzyme

expression levels. However, protein abundance would still have to be assessed, as protein stability and turnover rate will not be equalised by IntF2A.

Another strategy to improve metabolic channelling in the compartments is by not only using probabilistic metabolic channelling (i.e.: the high local concentration of subsequent enzymes in a metabolic pathway), but direct metabolic channelling as well. It was previously shown that directly linking enzymes via shared hydration layers (e.g., short protein linkers) can elicit metabolic channelling as it can help elicit substrate retention[251]. This could be achieved for a two enzyme microdomain by genetic fusion, or the attachment of an alternative, orthogonal SpyTag system (such as SnoopTag)[252], so that two different cargos with subsequent enzymes in a pathway are not only attached to the surface of the compartment (via the SpyCatcher – SpyTag system), but also to eachother via the orthogonal SnoopCatcher – SnoopTag system. Hetero-oligomeric scaffold constructs (mentioned above in 6.2.2.) could also be used for this purpose, with a different enzyme genetically fused for the different heteromers.

6.3.3 Transitioning to a stably expressing host to improve all putative applications

One of the most important improvements for the use of the compartment system for any applications is to transition to a stably expressing host. Transient expression via *A. tumefaciens*-mediated transformation is an effective method for rapid testing of new constructs and proof-of-principle experiments, but it has several drawbacks. Most importantly, protein expression levels are highly variable between cells due to the random integration of the genetic constructs into the host genome[253]. Furthermore, a substantial population of the cells are not transformed, which becomes a significant problem when

multiple constructs are co-transformed as the proportion of cells expressing all constructs decreases exponentially with the number of constructs used. For large-scale commercial uses these two major issues must be rectified.

All of these issues can be solved by developing and characterising an *N. benthamiana* line which stably expresses the compartment-forming construct C22Y-SpyT-his, which has both of the functionalising protein tags (SpyTag and 6xHisTag). This would remove a significant complexity from all experiments (since transforming plants with the compartment will no longer be necessary) and mean that all cells will express compartments, improving production efficicency. Importantly, 22Y and C22Y compartment phenotypes were identical in all visualised tissue types of stably transformed *A. thaliana* and transiently transformed mature leaves of *N. benthamiana* and *N. tabacum* (Figure 5.12). This is an important finding as it showcases that the compartment phenotype is consistent across different species, tissues and transformation types, increasing its potential utility and showing that the use of stably transformed hosts is likely a viable strategy. As such, developing this new host line should be top priority for any future work into the functionalisation of the compartment system.

6.3.4 Potential avenues for further long-term improvements to the compartment system for novel applications

Expression of high-value proteins and use as a metabolic microdomain were the focus of the functionalisation of the compartment system, but further applications could be envisioned (Figure 6.1).

It was discovered early on in the course of this project, that the compartment retains connections to the ER (Figure 3.9). However, the complete removal of the compartment

from the ER could be beneficial, as it would produce an independent synthetic membranous organelle. Although the specific removal of the compartment from the ER is currently beyond the capacity of the synthetic biology field, further basic and applied research into ER budding and fission of vesicles from it could reveal potential solutions to this challenge. One such putative line of inquiry is into the dynamin family of proteins, which were shown to be recruited to polymerise at high-curvature ER membrane tips and induce fission[254]. If a specific dynamin-recruiting signal would be discovered or engineered, it might be feasible to use this class of proteins to cleave the compartment off the ER (Figure 6.1A). A potential alternative approach could be to use Zera, a synthetic peptide derived from the plant storage protein γ-Zein, which was recently shown to be able to derive small membranous vesicles from the ER[77].

Completely separating the compartment from the ER would open up the possibilities for new functionalisations. Introduction of specific channel and pore molecules to the membrane of the compartment would allow the accumulation and storage of valuable biochemicals inside (Figure 6.1A). Pilot experiments could utilise AtSWEET11, a sucrose and glucose transporter[255], to store sugars already abundant in the plant cell. However, concentration and recovery of stored compounds would require a uniporter or a controllable channel (and would subsequently neccessitate a proton pump to energise the membrane) to ascertain that during purification the stored compounds are not lost. Commercially more interesting applications could include the concentration and purification of carboxylic acids used as industrial raw materials, such as fumarate and succinate[22,256,257]. Sequestration of valuable molecules can improve bioproduction by reducing potential negative feedback on the pathway as well, which was suggested as a

reason why the production of terpenoids on the surface of an OSER-like structure in *N. benthamiana* was so efficient in a recent study[258].

Once import of specific molecules into the compartment is achieved, intra-compartment enzyme pathways can be built as well, with an improved sequestration from the bulk of the cell, and products would be immediately compartmentalised, aiding purification (Figure 6.1A).

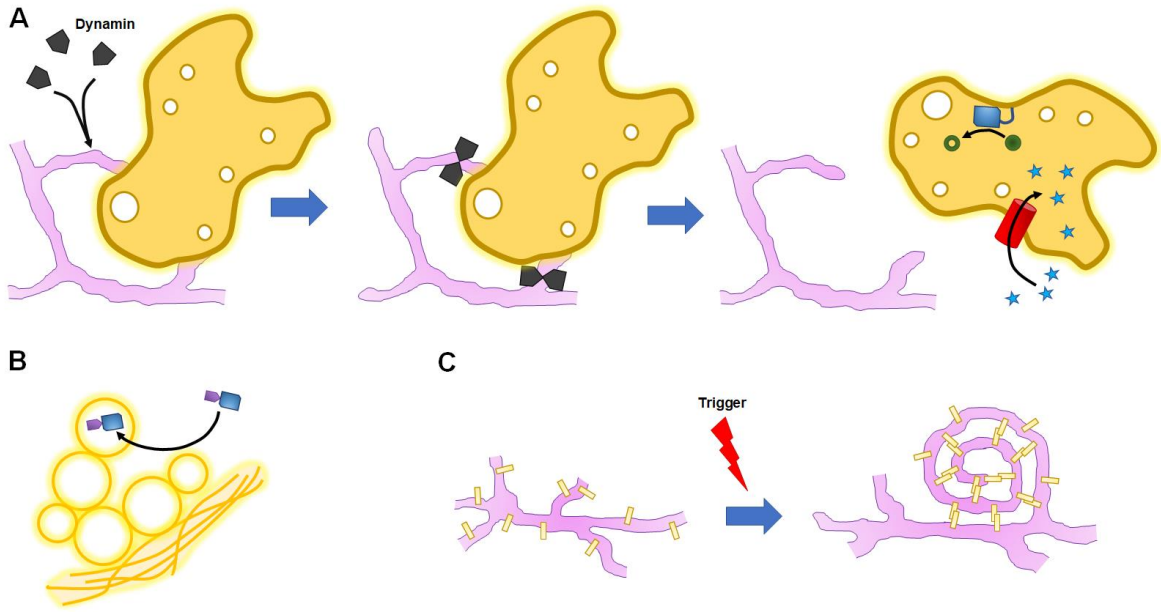


Figure 6.1: Potential improvements to the compartment system to enable further applications. A: Release of the compartment from the ER (for example via targeted dynamin association) would create a completely stand-alone synthetic membranous organelle. This would then enable of storage and accumulation of valuable biochemicals (blue stars) inside if appropriate channels or pores can be expressed to the compartment membrane (red tube). Intra-compartment enzymes (blue box) can build upon this by facilitating metabolic conversion of target molecules inside (green circles). B: Swollen ER bodies could be used for specific storage of HDEL-tagged proteins. C: Inducible scaffold molecules would enable triggerable formation of compartments.

The swollen ER bodies described in Chapter 3.2.4 (Figure 3.11) could also hold capacity for functionalisation. Since RFP-HDEL markers localised efficiently into the lumen of the swollen ER bodies generated by 22Y expression, it could be possible to use HDEL tags to sequester proteins into these vesicles (Figure 6.1B). However, further research into the formation of the swollen ER bodies is required since they were not always observed when expressing 22Y.

Another area of interest for long-term improvements to the compartment system could be the development of triggerable formation (Figure 6.1C). Many potential techniques can be investigated, such as recently popularised optogenetic approaches, with inducible formation of membraneless organelles now possible[64]. If this method could be extended to my system, it could provide a method for at-will formation of the compartments. Chemical inducers are also a possibility, as it was demonstrated in HeLa cells, where F_v domain dimerisation was triggered using the introduction of a chemical inducer AP20187, causing the rapid formation of OSER structures[139].

6.4 Comparison of the compartment system to the currently available methods of compartmentalisation

The compartment system has been investigated thoroughly during the course of this project. But it is important to put these results into context and compare them to the literature. Compartmentalisation is one of the key challenges of the synthetic biology field, and three distinct competing approaches have emerged to tackle it.

Co-opting existing organelles by modulating the localisation of recombinant proteins is commonly used for a variety of native organelles, such as peroxisomes[44,46,249,259], chloroplasts[57], mitochondria[53,54] or even the ER[58,247,248]. This method is less difficult in practice than others, but can negatively affect the host organelle, and endogenous pathways might compete for substrates when trying to reconstitutite metabolic pathways.

Building a novel membranous compartment would reduce crosstalk with host background, but building such a compartment and accurately targeting to it is very challenging, although there are recent promising experimental results using ER-derived synthetic membranous protein bodies in yeast[77]. Flavivirus-induced ER-derived membranous vesicles were also suggested as potential targets for functionalisation[79].

Building smaller membraneless organelles *in vivo* using phase-separation has become the popular third choice for the generation of compartments[61,64,66]. Virus-like particles can also be used for building protein-based compartments, and encapsualting metabolic pathways is possible into these, but these are very limited in size and capacity and can be harmful to the host cells[71,72].

Compared to these methods, the compartment system described above is inbetween targeting existing compartments and building a new membraneous compartment. It is delimited by a membrane bilayer, which is a clear benefit over membraneless organelles for both targeting membrane proteins to it, and being able to sequester compounds away from the bulk of the cell better. It is not completely standalone, as some connections to the ER are retained, but since there is a diffusion barrier to both the bulk of the cell and the rest of the ER, it does share some of the benefits of novel synthetic organelles. While it originates from the ER, and shares most of its properties with it, this system is far more advanced than a simple targeting of proteins to the bulk of the ER, and its impacts on the host organelle are very limited based on phenotypic analysis. Finally, a single scaffold construct is sufficient to elicit the formation of the compartment, similar to most phase-separation based membraneless organelles, which are preferred for their simplicity.

Overall, based on these features, the compartment system originally described by Ian Moore and Marketa Samalova which was characterised in-depth in this thesis undoubtedly shares many desirable properties with the currently available state-of-the-art compartmentalisation strategies. As such, it is definitely a worthwhile contestant to become a popular synthetic biology tool for plants.

References

1. Schillberg, S. & Finnern, R. Plant molecular farming for the production of valuable proteins - Critical evaluation of achievements and future challenges. *J. Plant Physiol.* **258–259**, 153359 (2021).
2. Stieger, M., Neuhaus, G., Momma, T., Schell, J. & Kreuzaler, F. Self assembly of immunoglobulins in the cytoplasm of the alga *Acetabularia mediterranea*. *Plant Sci.* **73**, 181–190 (1991).
3. Fischer, R. & Buyel, J. F. Molecular farming - The slope of enlightenment. *Biotechnol. Adv.* **40**, 107519 (2020).
4. Ma, J. K.-C. *et al.* Regulatory approval and a first-in-human phase I clinical trial of a monoclonal antibody produced in transgenic tobacco plants. *Plant Biotechnol. J.* **13**, 1106–1120 (2015).
5. Topp, E. *et al.* The case for plant-made veterinary immunotherapeutics. *Biotechnol. Adv.* **34**, 597–604 (2016).
6. Kurup, V. M. & Thomas, J. Edible vaccines: promises and challenges. *Mol. Biotechnol.* **62**, 79–90 (2020).
7. Reski, R., Bae, H. & Simonsen, H. T. *Physcomitrella patens*, a versatile synthetic biology chassis. *Plant Cell Rep.* **37**, 1409–1417 (2018).
8. Zimran, A., Wajnrajch, M., Hernandez, B. & Pastores, G. M. Taliglucerase alfa: safety and efficacy across 6 clinical studies in adults and children with Gaucher disease. *Orphanet J. Rare Dis.* **13**, 36 (2018).
9. Schillberg, S., Raven, N., Spiegel, H., Rasche, S. & Buntru, M. Critical analysis of the commercial potential of plants for the production of recombinant proteins. *Front. Plant Sci.* **10**, 720 (2019).
10. Mirzaee, M., Osmani, Z., Frébortová, J. & Frébort, I. Recent advances in molecular farming using monocot plants. *Biotechnol. Adv.* 107913 (2022).

11. Yang, J., Barr, L. A., Fahnestock, S. R. & Liu, Z.-B. High yield recombinant silk-like protein production in transgenic plants through protein targeting. *Transgenic Res.* **14**, 313–324 (2005).
12. Tschofen, M., Knopp, D., Hood, E. & Stöger, E. Plant molecular farming: much more than medicines. *Annu. Rev. Anal. Chem.* **9**, 271–294 (2016).
13. Moses, T., Mehrshahi, P., Smith, A. G. & Goossens, A. Synthetic biology approaches for the production of plant metabolites in unicellular organisms. *J. Exp. Bot.* **68**, 4057–4074 (2017).
14. O'Neill, E. C. & Kelly, S. Engineering biosynthesis of high-value compounds in photosynthetic organisms. *Crit. Rev. Biotechnol.* **37**, 779–802 (2017).
15. Ikram, N. K. B. K. & Simonsen, H. T. A Review of biotechnological artemisinin production in plants. *Front. Plant Sci.* **8**, 1966 (2017).
16. Exposito, O. *et al.* Metabolic responses of *Taxus media* transformed cell cultures to the addition of methyl jasmonate. *Biotechnol. Prog.* **26**, 1145–1153 (2010).
17. Matsuura, H. N. *et al.* Specialized plant metabolism characteristics and impact on target molecule biotechnological production. *Mol. Biotechnol.* **60**, 169–183 (2018).
18. Reed, J. & Osbourn, A. Engineering terpenoid production through transient expression in *Nicotiana benthamiana*. *Plant Cell Rep.* **37**, 1431–1441 (2018).
19. Lee, J. W. Synthetic biology: a new opportunity in the field of plant biochemistry & physiology. *J Plant Biochem Physiol* **1**, 1–2 (2013).
20. Fesenko, E. & Edwards, R. Plant synthetic biology: a new platform for industrial biotechnology. *J. Exp. Bot.* **65**, 1927–1937 (2014).
21. Buyel, J. F., Twyman, R. M. & Fischer, R. Very-large-scale production of antibodies in plants: The biologization of manufacturing. *Biotechnol. Adv.* **35**, 458–465 (2017).
22. Ahmad, P. *et al.* Role of transgenic plants in agriculture and biopharming. *Biotechnol. Adv.* **30**, 524–540 (2012).
23. Rademacher, T. *et al.* Plant cell packs: a scalable platform for recombinant protein production and metabolic engineering. *Plant Biotechnol. J.* **17**, 1560–1566 (2019).

24. Karki, U., Fang, H., Guo, W., Unnold-Cofre, C. & Xu, J. Cellular engineering of plant cells for improved therapeutic protein production. *Plant Cell Rep.* **40**, 1087–1099 (2021).

25. Buntru, M., Vogel, S., Spiegel, H. & Schillberg, S. Tobacco BY-2 cell-free lysate: an alternative and highly-productive plant-based *in vitro* translation system. *BMC Biotechnol.* **14**, 37 (2014).

26. Holtz, B. R. *et al.* Commercial-scale biotherapeutics manufacturing facility for plant-made pharmaceuticals. *Plant Biotechnol. J.* **13**, 1180–1190 (2015).

27. Xu, J., Towler, M. & Weathers, P. J. Platforms for plant-based protein production. in *Bioprocessing of Plant In Vitro Systems* (ed. Bley, A. P.) 509–548 (Nature Publishing Group, 2018).

28. Dhama, K. *et al.* Plant-based vaccines and antibodies to combat COVID-19: current status and prospects. *Hum. Vaccin. Immunother.* **16**, 2913–2920 (2020).

29. Wada, N., Ueta, R., Osakabe, Y. & Osakabe, K. Precision genome editing in plants: state-of-the-art in CRISPR/Cas9-based genome engineering. *BMC Plant Biol.* **20**, 234 (2020).

30. Fischer, R., Stoger, E., Schillberg, S., Christou, P. & Twyman, R. M. Plant-based production of biopharmaceuticals. *Curr. Opin. Plant Biol.* **7**, 152–158 (2004).

31. Polka, J. K., Hays, S. G. & Silver, P. A. Building spatial synthetic biology with compartments, scaffolds, and communities. *Cold Spring Harb. Perspect. Biol.* **8**, (2016).

32. Bolsover, S. R., Hyams, J. S., Shephard, E. A. & White, H. A. Membranes and organelles. in *Cell Biology: A Short Course* (ed. Bolsover, S. R., Hyams, J. S., Shephard, E. A. & White, H. A) 43–51 (John Wiley & Sons, 2011).

33. Monnard, P.-A. & Walde, P. Current ideas about prebiological compartmentalization. *Life* **5**, 1239–1263 (2015).

34. Kleckner, N. *et al.* The bacterial nucleoid: nature, dynamics and sister segregation. *Curr. Opin. Microbiol.* **22**, 127–137 (2014).

35. Agapakis, C. M., Boyle, P. M. & Silver, P. A. Natural strategies for the spatial optimization of metabolism in synthetic biology. *Nat. Chem. Biol.* **8**, 527–535 (2012).

36. Kerfeld, C. A., Aussignargues, C., Zarzycki, J., Cai, F. & Sutter, M. Bacterial microcompartments. *Nat. Rev. Microbiol.* **16**, 277–290 (2018).

37. Mao, Y. S., Zhang, B. & Spector, D. L. Biogenesis and function of nuclear bodies. *Trends Genet.* **27**, 295–306 (2011).

38. Kaiser, T. E., Intine, R. V. & Dundr, M. *De novo* formation of a subnuclear body. *Science* **322**, 1713–1717 (2008).

39. Cuevas-Velazquez, C. L. & Dinneny, J. R. Organization out of disorder: liquid-liquid phase separation in plants. *Curr. Opin. Plant Biol.* **45**, 68–74 (2018).

40. Boeynaems, S. *et al.* Protein phase separation: A new phase in cell biology. *Trends Cell Biol.* **28**, 420–435 (2018).

41. Sweetlove, L. J. & Fernie, A. R. The spatial organization of metabolism within the plant cell. *Annu. Rev. Plant Biol.* **64**, 723–746 (2013).

42. Cárdenas, P. D., Almeida, A. & Bak, S. Evolution of structural diversity of triterpenoids. *Front. Plant Sci.* **10**, 1523 (2019).

43. De Meyer, T. & Depicker, A. Trafficking of endoplasmic reticulum-retained recombinant proteins is unpredictable in *Arabidopsis thaliana*. *Front. Plant Sci.* **5**, 473 (2014).

44. Gao, J. & Zhou, Y. J. Repurposing peroxisomes for microbial synthesis for biomolecules. *Methods Enzymol.* **617**, 83–111 (2019).

45. Zhou, Y. J. *et al.* Harnessing yeast peroxisomes for biosynthesis of fatty-acid-derived biofuels and chemicals with relieved side-pathway competition. *J. Am. Chem. Soc.* **138**, 15368–15377 (2016).

46. Herr, A. & Fischer, R. Improvement of *Aspergillus nidulans* penicillin production by targeting AcvA to peroxisomes. *Metab. Eng.* **25**, 131–139 (2014).

47. DeLoache, W. C., Russ, Z. N. & Dueber, J. E. Towards repurposing the yeast peroxisome for compartmentalizing heterologous metabolic pathways. *Nat. Commun.* **7**, 11152 (2016).

48. Thodey, K., Galanie, S. & Smolke, C. D. A microbial biomanufacturing platform for natural and semisynthetic opioids. *Nat. Chem. Biol.* **10**, 837–844 (2014).

49. Bayer, T. S. *et al.* Synthesis of methyl halides from biomass using engineered microbes. *J. Am. Chem. Soc.* **131**, 6508–6515 (2009).

50. Fan, L.-H., Zhang, Z.-J., Yu, X.-Y., Xue, Y.-X. & Tan, T.-W. Self-surface assembly of cellulosomes with two miniscaffoldins on *Saccharomyces cerevisiae* for cellulosic ethanol production. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 13260–13265 (2012).

51. Avalos, J. L., Fink, G. R. & Stephanopoulos, G. Compartmentalization of metabolic pathways in yeast mitochondria improves the production of branched-chain alcohols. *Nat. Biotechnol.* **31**, 335–341 (2013).

52. Hammer, S. K. & Avalos, J. L. Harnessing yeast organelles for metabolic engineering. *Nat. Chem. Biol.* **13**, 823–832 (2017).

53. Liu, Q. *et al.* Reconstitution of the costunolide biosynthetic pathway in yeast and *Nicotiana benthamiana*. *PLoS One* **6**, e23255 (2011).

54. Eljounaidi, K. *et al.* Cytochrome P450s from *Cynara cardunculus* L. CYP71AV9 and CYP71BL5, catalyze distinct hydroxylations in the sesquiterpene lactone biosynthetic pathway. *Plant Sci.* **223**, 59–68 (2014).

55. van Herpen, T. W. J. M. *et al. Nicotiana benthamiana* as a production platform for artemisinin precursors. *PLoS One* **5**, e14222 (2010).

56. Henriques de Jesus, M. P. R. *et al.* Tat proteins as novel thylakoid membrane anchors organize a biosynthetic pathway in chloroplasts and increase product yield 5-fold. *Metab. Eng.* **44**, 108–116 (2017).

57. Malhotra, K. *et al.* Compartmentalized metabolic engineering for artemisinin biosynthesis and effective malaria treatment by oral delivery of plant cells. *Mol. Plant* **9**, 1464–1477 (2016).

58. Ding, B.-J. *et al.* A plant factory for moth pheromone production. *Nat. Commun.* **5**, 3353 (2014).

59. Alberti, S. Phase separation in biology. *Curr. Biol.* **27**, R1097–R1102 (2017).

60. Uversky, V. N. Supramolecular fuzziness of intracellular liquid droplets: Liquid-liquid phase transitions, membrane-less organelles, and intrinsic disorder. *Molecules* **24**, (2019).

61. Bracha, D., Walls, M. T. & Brangwynne, C. P. Probing and engineering liquid-phase organelles. *Nat. Biotechnol.* **37**, 1435–1445 (2019).

62. Nott, T. J. *et al.* Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell* **57**, 936–947 (2015).

63. Schuster, B. S. *et al.* Controllable protein phase separation and modular recruitment to form responsive membraneless organelles. *Nat. Commun.* **9**, 2985 (2018).

64. Zhao, E. M. *et al.* Light-based control of metabolic flux through assembly of synthetic organelles. *Nat. Chem. Biol.* **15**, 589–597 (2019).

65. Banani, S. F. *et al.* Compositional control of phase-separated cellular bodies. *Cell* **166**, 651–663 (2016).

66. Garabedian, M. V. *et al.* Designer membraneless organelles sequester native factors for control of cell behavior. *Nat. Chem. Biol.* **17**, 998–1007 (2021).

67. Reinkemeier, C. D., Girona, G. E. & Lemke, E. A. Designer membraneless organelles enable codon reassignment of selected mRNAs in eukaryotes. *Science* **363**, eaaw2644 (2019).

68. Mackinder, L. C. M. The *Chlamydomonas* CO2 -concentrating mechanism and its potential for engineering photosynthesis in plants. *New Phytol.* **217**, 54–61 (2018).

69. Smirnoff, N. Engineering of metabolic pathways using synthetic enzyme complexes. *Plant Physiol.* **179**, 918–928 (2019).

70. Giessen, T. B. W. & Silver, P. A. A catalytic nanoreactor based on *in vivo* encapsulation of multiple enzymes in an engineered protein nanocompartment. *ChemBioChem* 1931–1935 (2016).

71. Cheah, L. C. *et al.* Artificial self-assembling nanocompartment for organizing metabolic pathways in yeast. *ACS Synth. Biol.* **10**, 3251–3263 (2021).

72. Catrice, E. V. B. & Sainsbury, F. Assembly and purification of polyomavirus-like particles from plants. *Mol. Biotechnol.* **57**, 904–913 (2015).

73. Fan, C., Cheng, S., Sinha, S. & Bobik, T. A. Interactions between the termini of lumen enzymes and shell proteins mediate enzyme encapsulation into bacterial microcompartments. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 14995–15000 (2012).

74. Lau, Y. H., Giessen, T. W., Altenburg, W. J. & Silver, P. A. Prokaryotic nanocompartments form synthetic organelles in a eukaryote. *Nat. Commun.* **9**, 1311 (2018).

75. Lin, M. T. *et al.* β-Carboxysomal proteins assemble into highly organized structures in *Nicotiana chloroplasts*. *Plant J.* **79**, 1–12 (2014).

76. Long, B. M. *et al.* Carboxysome encapsulation of the CO2-fixing enzyme Rubisco in tobacco chloroplasts. *Nat. Commun.* **9**, 3570 (2018).

77. Reifenrath, M., Oreb, M., Boles, E. & Tripp, J. Artificial ER-derived vesicles as synthetic organelles for *in vivo* compartmentalization of biochemical pathways. *ACS Synth. Biol.* **9**, 2909–2916 (2020).

78. Wu, T. *et al.* Engineering an artificial membrane vesicle trafficking system (AMVTS) for the excretion of β-carotene in *Escherichia coli*. *ACS Synth. Biol.* **8**, 1037–1046 (2019).

79. Arakawa, M. & Morita, E. Flavivirus replication organelle biogenesis in the endoplasmic reticulum: comparison with other single-stranded positive-sense RNA viruses. *Int. J. Mol. Sci.* **20**, (2019).

80. Elani, Y., Law, R. V. & Ces, O. Vesicle-based artificial cells as chemical microreactors with spatially segregated reaction pathways. *Nat. Commun.* **5**, 5305 (2014).

81. Smith, J. M., Chowdhry, R. & Booth, M. J. Controlling synthetic cell-cell communication. *Front Mol Biosci* **8**, 809945 (2021).

82. Wheeldon, I. *et al.* Substrate channelling as an approach to cascade reactions. *Nat. Chem.* **8**, 299–309 (2016).

83. Sweetlove, L. J. & Fernie, A. R. The role of dynamic enzyme assemblies and substrate channelling in metabolic regulation. *Nat. Commun.* **9**, 2136 (2018).

84. Wu, F. & Minteer, S. Krebs cycle metabolon: structural evidence of substrate channeling revealed by cross-linking and mass spectrometry. *Angew. Chem. Int. Ed Engl.* **54**, 1851–1854 (2015).

85. Bulutoglu, B., Garcia, K. E., Wu, F., Minteer, S. D. & Banta, S. Direct evidence for metabolon formation and substrate channeling in recombinant TCA cycle enzymes. *ACS Chem. Biol.* **11**, 2847–2853 (2016).

86. Zhang, Y. *et al.* Protein-protein interactions and metabolite channelling in the plant tricarboxylic acid cycle. *Nat. Commun.* **8**, 15212 (2017).

87. Fernie, A. R., Zhang, Y. & Sweetlove, L. J. Passing the baton: Substrate channelling in respiratory metabolism. *Research* **2018**, 1539325 (2018).

88. Laursen, T. *et al.* Characterization of a dynamic metabolon producing the defense compound dhurrin in sorghum. *Science* **354**, 890–893 (2016).

89. Pareek, V., Tian, H., Winograd, N. & Benkovic, S. J. Metabolomics and mass spectrometry imaging reveal channeled *de novo* purine synthesis in cells. *Science* **368**, 283–290 (2020).

90. Gou, M., Ran, X., Martin, D. W. & Liu, C.-J. The scaffold proteins of lignin biosynthetic cytochrome P450 enzymes. *Nat Plants* **4**, 299–310 (2018).

91. Zhang, Y. & Fernie, A. R. Metabolons, enzyme-enzyme assemblies that mediate substrate channeling, and their roles in plant metabolism. *Plant Commun* **2**, 100081 (2021).

92. Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W. & Davies, D. R. Three-dimensional structure of the tryptophan synthase alpha 2 beta 2 multienzyme complex from *Salmonella typhimurium*. *J. Biol. Chem.* **263**, 17857–17871 (1988).

93. Anderson, K. S. *et al.* Kinetic characterization of channel impaired mutants of tryptophan synthase (∗). *J. Biol. Chem.* **270**, 29936–29944 (1995).

94. Weeks, A., Lund, L. & Raushel, F. M. Tunneling of intermediates in enzyme-catalyzed reactions. *Curr. Opin. Chem. Biol.* **10**, 465–472 (2006).

95. Perham, R. N. Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions. *Annu. Rev. Biochem.* **69**, 961–1004 (2000).

96. Smolle, M. & Lindsay, J. G. Molecular architecture of the pyruvate dehydrogenase complex: bridging the gap. *Biochem. Soc. Trans.* **34**, 815–818 (2006).

97. Vickery, C. R., La Clair, J. J., Burkart, M. D. & Noel, J. P. Harvesting the biosynthetic machineries that cultivate a variety of indispensable plant natural products. *Curr. Opin. Chem. Biol.* **31**, 66–73 (2016).

98. Castellana, M. *et al.* Enzyme clustering accelerates processing of intermediates through metabolic channeling. *Nat. Biotechnol.* **32**, 1011–1018 (2014).

99. Lee, H., DeLoache, W. C. & Dueber, J. E. Spatial organization of enzymes for metabolic engineering. *Metab. Eng.* **14**, 242–251 (2012).

100. Alam, M. T. *et al.* The self-inhibitory nature of metabolic networks and its alleviation through compartmentalization. *Nat. Commun.* **8**, 16018 (2017).

101. Poshyvailo, L., von Lieres, E. & Kondrat, S. Does metabolite channeling accelerate enzyme-catalyzed cascade reactions? *PLoS One* **12**, e0172673 (2017).

102. Fu, J. *et al.* Multi-enzyme complexes on DNA scaffolds capable of substrate channelling with an artificial swinging arm. *Nat. Nanotechnol.* **9**, 531–536 (2014).

103. Liu, Y. *et al.* Substrate Channeling in an artificial metabolon: A molecular dynamics blueprint for an experimental peptide bridge. *ACS Catal.* **7**, 2486–2493 (2017).

104. Pareek, V., Sha, Z., He, J., Wingreen, N. S. & Benkovic, S. J. Metabolic channeling: predictions, deductions, and evidence. *Mol. Cell* **81**, 3775–3785 (2021).

105. Obata, T. Metabolons in plant primary and secondary metabolism. *Phytochem. Rev.* **18**, 1483–1507 (2019).

106. Almsherqi, Z. A., Landh, T., Kohlwein, S. D. & Deng, Y. Cubic membranes: the missing dimension of cell membrane organization. *Int. Rev. Cell Mol. Biol.* **274**, 275–342 (2009).

107. Stefano, G. & Brandizzi, F. Advances in plant ER architecture and dynamics. *Plant Physiol.* **176**, 178–186 (2018).

108. Sparkes, I., Hawes, C. & Frigerio, L. FrontiERs: movers and shapers of the higher plant cortical endoplasmic reticulum. *Curr. Opin. Plant Biol.* **14**, 658–665 (2011).

109. Pain, C. & Kriechbaumer, V. Defining the dance: quantification and classification of endoplasmic reticulum dynamics. *J. Exp. Bot.* **71**, 1757–1762 (2020).

110. Voeltz, G. K., Prinz, W. A., Shibata, Y., Rist, J. M. & Rapoport, T. A. A class of membrane proteins shaping the tubular endoplasmic reticulum. *Cell* **124**, 573–586 (2006).

111. Shibata, Y. *et al.* The reticulon and DP1/Yop1p proteins form immobile oligomers in the tubular endoplasmic reticulum. *J. Biol. Chem.* **283**, 18892–18904 (2008).

112. Sparkes, I. *et al.* Five *Arabidopsis* reticulon isoforms share endoplasmic reticulum location, topology, and membrane-shaping properties. *Plant Cell* **22**, 1333–1343 (2010).

113. Kriechbaumer, V. *et al. Arabidopsis* Lunapark proteins are involved in ER cisternae formation. *New Phytol.* **219**, 990–1004 (2018).

114. Chen, J., Stefano, G., Brandizzi, F. & Zheng, H. *Arabidopsis* RHD3 mediates the generation of the tubular ER network and is required for Golgi distribution and motility in plant cells. *J. Cell Sci.* **124**, 2241–2252 (2011).

115. Ueda, H. *et al.* Phosphorylation of the C Terminus of RHD3 Has a Critical Role in Homotypic ER Membrane Fusion in *Arabidopsis*. *Plant Physiol.* **170**, 867–880 (2016).

116. Shibata, Y., Voeltz, G. K. & Rapoport, T. A. Rough sheets and smooth tubules. *Cell* **126**, 435–439 (2006).

117. Snapp, E. L. *et al.* Formation of stacked ER cisternae by low affinity protein interactions. *J. Cell Biol.* **163**, 257–269 (2003).

118. Borgese, N., Francolini, M. & Snapp, E. Endoplasmic reticulum architecture: structures in flux. *Curr. Opin. Cell Biol.* **18**, 358–364 (2006).

119. Schäfer, J. A. *et al.* ESCRT machinery mediates selective microautophagy of endoplasmic reticulum in yeast. *EMBO J.* **39**, e102586 (2020).

120. Porter, K. R. & Yamada, E. Studies on the endoplasmic reticulum. Its form and differentiation in pigment epithelial cells of the frog retina. *J. Biophys. Biochem. Cytol.* **8**, 181–205 (1960).

121. Tabor, G. A. & Fisher, S. K. Myeloid bodies in the mammalian retinal pigment epithelium. *Invest. Ophthalmol. Vis. Sci.* **24**, 388–391 (1983).

122. Naftolin, F., Bruhlmann-Papazyan, M., Baetens, D. & Garcia-Segura, L. M. Neurons with whorl bodies have increased numbers of synapses. *Brain Res.* **329**, 289–293 (1985).

123. Takei, K., Mignery, G. A., Mugnaini, E., Südhof, T. C. & De Camilli, P. Inositol 1,4,5-trisphosphate receptor causes formation of ER cisternal stacks in transfected fibroblasts and in cerebellar Purkinje cells. *Neuron* **12**, 327–342 (1994).

124. Esau, K. & Gill, R. H. Aggregation of endoplasmic reticulum and its relation to the nucleus in a differentiating sieve element. *J. Ultrastruct. Res.* **34**, 144–158 (1971).

125. Oparka, K. J. & Johnson, R. P. Endoplasmic reticulum and crystalline fibrils in the root protophloem of *Nymphoides peltata*. *Planta* **143**, 21–27 (1978).

126. Esau, K. & Hoefert, I. L. Endoplasmic reticulum and its relation to microtubules in sieve elements of sugarbeet and spinach. *J. Ultrastruct. Res.* **71**, 249–257 (1980).

127. Esau, K. & Cronshaw, J. Endoplasmic reticulum in the sieve element of *Cucurbita*. *J. Ultrastruct. Res.* **23**, 1–14 (1968).

128. Thorsch, J. & Esau, K. Changes in the endoplasmic reticulum during differentaiation of a sieve element in *Gossypium hirsutum*. *J. Ultrastruct. Res.* **74**, 183–194 (1981).

129. Parthasarathy, M. V. Ultrastructure of phloem in palms. *Protoplasma* **79**, 265–315 (1974).

130. Evert, R. F. & Deshpande, B. P. Electron microscope investigation of sieve-element ontogeny and structure in *Ulmus americana*. *Protoplasma* **68**, 403–432 (1969).

131. Wooding, F. B. P. Endoplasmic reticulum aggregates of ordered structure. *Planta* **76**, 205–208 (1967).

132. Behnke, H. D. Sieve–element characters. *Nord. J. Bot.* **1**, 381–400 (1981).

133. Sennerby-Forsse, L. & von Fircks, H. A. Ultrastructure of cells in the cambial region during winter hardening and spring dehardening in *Salix dasyclados Wim.* grown at two nutrient levels. *Trees* **1**, 151–163 (1987).

134. Sennerby-Forsse, L., Singh, A. P. & Walles, B. Presence of myelin-like structures in tanniferous vascular parenchyma cells of *Salix dasyclados* Wim. *Protoplasma* **138**, 183–186 (1987).

135. Cresti, M., Pacini, E., Ciampolini, F. & Sarfatti, G. Germination and early tube development in vitro of *Lycopersicum peruvianum* pollen: Ultrastructural features. *Planta* **136**, 239–247 (1977).

136. Cresti, M., Ciampolini, F., Mulcahy, D. L. M. & Mulcahy, G. Ultrastructure of *Nicotiana alata* pollen, its germination and early tube formation. *Am. J. Bot.* **72**, 719–727 (1985).

137. Yamamoto, A., Masaki, R. & Tashiro, Y. Formation of crystalloid endoplasmic reticulum in COS cells upon overexpression of microsomal aldehyde dehydrogenase by cDNA transfection. *J. Cell Sci.* **109**, 1727–1738 (1996).

138. Korkhov, V. M. *et al.* Peptide-based interactions with calnexin target misassembled membrane proteins into endoplasmic reticulum-derived multilamellar bodies. *J. Mol. Biol.* **378**, 337–352 (2008).

139. Lingwood, D., Schuck, S., Ferguson, C., Gerl, M. J. & Simons, K. Generation of cubic membranes by controlled homotypic interaction of membrane proteins in the endoplasmic reticulum. *J. Biol. Chem.* **284**, 12041–12048 (2009).

140. Volkova, E. G., Abramchuk, S. S. & Sheval, E. V. The overexpression of nuclear envelope protein Lap2β induces endoplasmic reticulum reorganisation via membrane stacking. *Biol. Open* **1**, 802–805 (2012).

141. Federovitch, C. M. *et al.* Genetic and structural analysis of Hmg2p-induced endoplasmic reticulum remodeling in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* **19**, 4506–4520 (2008).

142. Emmerstorfer, A. *et al.* Over-expression of ICE2 stabilizes cytochrome P450 reductase in *Saccharomyces cerevisiae* and *Pichia pastoris*. *Biotechnol. J.* **10**, 623–635 (2015).

143. Sandor, A., Fricker, M. D., Kriechbaumer, V. & Sweetlove, L. J. IntEResting structures: formation and applications of organized smooth endoplasmic reticulum in plant cells. *Plant Physiol.* **185**, 550–561 (2021).

144. Gong, F. C., Giddings, T. H., Meehl, J. B., Staehelin, L. A. & Galbraith, D. W. Z-membranes: artificial organelles for overexpressing recombinant integral membrane proteins. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 2219–2223 (1996).

145. Costantini, L. M., Fossati, M., Francolini, M. & Snapp, E. L. Assessing the tendency of fluorescent proteins to oligomerize under physiologic conditions. *Traffic* **13**, 643–649 (2012).

146. Fukuda, M., Yamamoto, A. & Mikoshiba, K. Formation of crystalloid endoplasmic reticulum induced by expression of synaptotagmin lacking the conserved WHXL motif in the C terminus. Structural importance of the WHXL motif in the C2B domain. *J. Biol. Chem.* **276**, 41112–41119 (2001).

147. Barbante, A. *et al.* Anchorage to the cytosolic face of the endoplasmic reticulum membrane: a new strategy to stabilize a cytosolic recombinant antigen in plants. *Plant Biotechnol. J.* **6**, 560–575 (2008).

148. Ferrero, S. *et al.* Proliferation and morphogenesis of the endoplasmic reticulum driven by the membrane domain of 3-hydroxy-3-methylglutaryl coenzyme A reductase in plant cells. *Plant Physiol.* **168**, 899–914 (2015).

149. Grados-Torrez, R. E., López-Iglesias, C., Ferrer, J. C. & Campos, N. Loose morphology and high dynamism of OSER structures induced by the membrane domain of HMG-CoA reductase. *Int. J. Mol. Sci.* **22**, 9132 (2021).

150. Marti, L. *et al.* A missense mutation in the vacuolar protein GOLD36 causes organizational defects in the ER and aberrant protein trafficking in the plant secretory pathway: GOLD36, a vacuolar protein for ER plant integrity. *Plant J.* **63**, 901–913 (2010).
151. Jancowski, S. *et al.* Trafficking of the myrosinase-associated protein GLL23 requires NUC/MVP1/GOLD36/ERMO3 and the p24 protein CYB. *Plant J.* **77**, 497–510 (2014).
152. Sprocati, T., Ronchi, P., Raimondi, A., Francolini, M. & Borgese, N. Dynamic and reversible restructuring of the ER induced by PDMP in cultured cells. *J. Cell Sci.* **119**, 3249–3260 (2006).
153. Paris, N. *et al.* Molecular cloning and further characterization of a probable plant vacuolar sorting receptor. *Plant Physiol.* **115**, 29–39 (1997).
154. Brandizzi, F. *et al.* The destination for single-pass membrane proteins is influenced markedly by the length of the hydrophobic domain. *Plant Cell* **14**, 1077–1092 (2002).
155. Meuli, N. The biogenesis of a novel endomembrane structure. (University of Oxford, 2005).
156. White, M. Biogenesis of a novel endomembrane storage compartment. (University of Oxford, 2006).
157. Kriechbaumer, V. & Brandizzi, F. The plant endoplasmic reticulum: an organized chaos of tubules and sheets with multiple functions. *J. Microsc.* **280**, 122–133 (2020).
158. Nelson, B. K., Cai, X. & Nebenführ, A. A multicolored set of *in vivo* organelle markers for co-localization studies in *Arabidopsis* and other plants. *Plant J.* **51**, 1126–1136 (2007).
159. Kriechbaumer, V., Botchway, S. W. & Hawes, C. Localization and interactions between *Arabidopsis* auxin biosynthetic enzymes in the TAA/YUC-dependent pathway. *J. Exp. Bot.* **67**, 4195–4207 (2016).
160. Sparkes, I. A., Hawes, C. & Baker, A. AtPEX2 and AtPEX10 are targeted to peroxisomes independently of known endoplasmic reticulum trafficking routes. *Plant Physiol.* **139**, 690–700 (2005).
161. Hung, Y. P. & Yellen, G. Live-cell imaging of cytosolic NADH-NAD+ redox state using a genetically encoded fluorescent biosensor. *Methods Mol. Biol.* **1071**, 83–95 (2014).

162. Keeble, A. H. *et al.* Approaching infinite affinity through engineering of peptide-protein interaction. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 26523–26533 (2019).

163. Karimi, M., Inzé, D. & Depicker, A. GATEWAY vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci.* **7**, 193–195 (2002).

164. Hanahan, D. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**, 557–580 (1983).

165. Zhao, H., Tan, Z., Wen, X. & Wang, Y. An improved syringe agroinfiltration protocol to enhance transformation efficiency by combinative use of 5-azacytidine, ascorbate acid and Tween-20. *Plants* **6**, (2017).

166. Voinnet, O., Rivas, S., Mestre, P. & Baulcombe, D. An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J.* **33**, 949–956 (2003).

167. Thévenaz, P., Ruttimann, U. E. & Unser, M. A pyramid approach to subpixel registration based on intensity. *IEEE Trans. Image Process.* **7**, 27–41 (1998).

168. Pain, C., Kriechbaumer, V., Kittelmann, M., Hawes, C. & Fricker, M. Quantitative analysis of plant ER architecture and dynamics. *Nat. Commun.* **10**, 984 (2019).

169. Kittelmann, M. 3D electron microscopy of the ER. in *The Plant Endoplasmic Reticulum : Methods and Protocols* (eds. Hawes, C. & Kriechbaumer, V.) 15–21 (Springer New York, 2018).

170. Lisec, J., Schauer, N., Kopka, J., Willmitzer, L. & Fernie, A. R. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat. Protoc.* **1**, 387–396 (2006).

171. Fletcher, J. M. *et al.* A basis set of de novo coiled-coil peptide oligomers for rational protein design and synthetic biology. *ACS Synth. Biol.* **1**, 240–250 (2012).

172. Gordon-Smith, D. J. *et al.* Solution structure of a C-terminal coiled-coil domain from bovine IF(1): the inhibitor protein of F(1) ATPase. *J. Mol. Biol.* **308**, 325–339 (2001).

173. Liu, Y. *et al.* The tetramer structure of the Nervy homology two domain, NHR2, is critical for AML1/ETO's activity. *Cancer Cell* **9**, 249–260 (2006).

174. Chen, X., Zaro, J. L. & Shen, W.-C. Fusion protein linkers: property, design and functionality. *Adv. Drug Deliv. Rev.* **65**, 1357–1369 (2013).

175. Bulbarelli, A., Sprocati, T., Barberi, M., Pedrazzini, E. & Borgese, N. Trafficking of tail-anchored proteins: transport from the endoplasmic reticulum to the plasma membrane and sorting between surface domains in polarised epithelial cells. *J. Cell Sci.* **115**, 1689–1702 (2002).

176. Sugimoto, N., Engelgau, P., Jones, A. D., Song, J. & Beaudry, R. Citramalate synthase yields a biosynthetic pathway for isoleucine and straight- and branched-chain ester formation in ripening apple fruit. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2009988118 (2021).

177. Ormö, M. *et al.* Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392–1395 (1996).

178. Zacharias, D. A., Violin, J. D., Newton, A. C. & Tsien, R. Y. Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* **296**, 913–916 (2002).

179. Au, K. K. C. *et al.* A perturbation in glutathione biosynthesis disrupts endoplasmic reticulum morphology and secretory membrane traffic in *Arabidopsis thaliana*. *Plant J.* **71**, 881–894 (2012).

180. Cranfill, P. J. *et al.* Quantitative assessment of fluorescent proteins. *Nat. Methods* **13**, 557–562 (2016).

181. Agee, A. E. *et al.* MODIFIED VACUOLE PHENOTYPE1 is an *Arabidopsis* myrosinase-associated protein involved in endomembrane protein trafficking. *Plant Physiol.* **152**, 120–132 (2010).

182. Nakano, R. T. *et al.* ERMO3/MVP1/GOLD36 is involved in a cell type-specific mechanism for maintaining ER morphology in *Arabidopsis thaliana*. *PLoS One* **7**, e49103 (2012).

183. Liu, H. F., Ma, J., Winter, C. & Bayer, R. Recovery and purification process development for monoclonal antibody production. *MAbs* **2**, 480–499 (2010).

184. Zakeri, B. *et al.* Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E690-7 (2012).

185. Reddington, S. C. & Howarth, M. Secrets of a covalent interaction for biomaterials and biotechnology: SpyTag and SpyCatcher. *Curr. Opin. Chem. Biol.* **29**, 94–99 (2015).

186. Bornhorst, J. A. & Falke, J. J. Purification of proteins using polyhistidine affinity tags. *Methods Enzymol.* **326**, 245–254 (2000).

187. Colas des Francs-Small, C., Ambard-Bretteville, F., Small, I. D. & Rémy, R. Identification of a major soluble protein in mitochondria from nonphotosynthetic tissues as NAD-dependent formate dehydrogenase. *Plant Physiol.* **102**, 1171–1177 (1993).

188. Rottensteiner, H. *et al.* Peroxisomal membrane proteins contain common Pex19p-binding sites that are an integral part of their targeting signals. *Mol. Biol. Cell* **15**, 3406–3417 (2004).

189. Kim, P. K. & Hettema, E. H. Multiple pathways for protein transport to peroxisomes. *J. Mol. Biol.* **427**, 1176–1190 (2015).

190. Cross, L. L., Ebeed, H. T. & Baker, A. Peroxisome biogenesis, protein targeting mechanisms and PEX gene functions in plants. *Biochim. Biophys. Acta* **1863**, 850–862 (2016).

191. Reumann, S. & Chowdhary, G. Prediction of peroxisomal matrix proteins in plants. in *Proteomics of Peroxisomes: Identifying Novel Functions and Regulatory Networks* (eds. del Río, L. A. & Schrader, M.) 125–138 (Springer Singapore, 2018).

192. Hoepfner, D., Schildknegt, D., Braakman, I., Philippsen, P. & Tabak, H. F. Contribution of the endoplasmic reticulum to peroxisome formation. *Cell* **122**, 85–95 (2005).

193. Stoddard, A. & Rolland, V. I see the light! Fluorescent proteins suitable for cell wall/apoplast targeting in *Nicotiana benthamiana* leaves. *Plant Direct* **3**, e00112 (2019).

194. Qu, S. *et al.* Ultramicroscopy reveals a layer of multiply folded membranes around the tannin-accumulating vacuole in honeysuckle petal trichomes. *Micron* **99**, 1–8 (2017).

195. Geertsma, E. R., Groeneveld, M., Slotboom, D.-J. & Poolman, B. Quality control of overexpressed membrane proteins. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 5722–5727 (2008).

196. Müller-Lucks, A., Bock, S., Wu, B. & Beitz, E. Fluorescent *in situ* folding control for rapid optimization of cell-free membrane protein synthesis. *PLoS One* **7**, e42186 (2012).

197. Bird, L. E. *et al.* Green fluorescent protein-based expression screening of membrane proteins in *Escherichia coli. J. Vis. Exp.* e52357 (2015).

198. Rath, A. & Deber, C. M. Correction factors for membrane protein molecular weight readouts on sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Anal. Biochem.* **434**, 67–72 (2013).

199. Kohler, H. P., Kohler-Staub, D. & Focht, D. D. Degradation of 2-hydroxybiphenyl and 2,2'-dihydroxybiphenyl by *Pseudomonas sp.* strain HBP1. *Appl. Environ. Microbiol.* **54**, 2683–2688 (1988).

200. Kohler, H. P., Schmid, A. & van der Maarel, M. Metabolism of 2,2'-dihydroxybiphenyl by *Pseudomonas sp.* strain HBP1: production and consumption of 2,2',3-trihydroxybiphenyl. *J. Bacteriol.* **175**, 1621–1628 (1993).

201. Suske, W. A. *et al.* Purification and characterization of 2-hydroxybiphenyl 3-monooxygenase, a novel NADH-dependent, FAD-containing aromatic hydroxylase from *Pseudomonas azelaica* HBP1. *J. Biol. Chem.* **272**, 24257–24265 (1997).

202. Odell, J. T., Nagy, F. & Chua, N. H. Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature* **313**, 810–812 (1985).

203. Pitzschke, A. *Agrobacterium* infection and plant defense-transformation success hangs by a thread. *Front. Plant Sci.* **4**, 519 (2013).

204. Le, X. H., Lee, C.-P., Monachello, D. & Harvey Millar, A. Metabolic evidence for distinct pyruvate pools inside plant mitochondria. *bioRxiv* 2021.10.30.466557 (2021) doi:10.1101/2021.10.30.466557.

205. Christianson, J. C. & Ye, Y. Cleaning up in the endoplasmic reticulum: ubiquitin in charge. *Nat. Struct. Mol. Biol.* **21**, 325–335 (2014).

206. Chen, Q., Yu, F. & Xie, Q. Insights into endoplasmic reticulum-associated degradation in plants. *New Phytol.* **226**, 345–350 (2020).

207. Gibbs, D. J., Bacardit, J., Bachmair, A. & Holdsworth, M. J. The eukaryotic N-end rule pathway: conserved mechanisms and diverse functions. *Trends Cell Biol.* **24**, 603–611 (2014).

208. Mot, A. C. *et al.* Real-time detection of N-end rule-mediated ubiquitination via fluorescently labeled substrate probes. *New Phytol.* **217**, 613–624 (2018).

209. Holdsworth, M. J., Vicente, J., Sharma, G., Abbas, M. & Zubrycka, A. The plant N-degron pathways of ubiquitin-mediated proteolysis. *J. Integr. Plant Biol.* **62**, 70–89 (2020).

210. Čapková, V., Hrabětová, E. & Tupý, J. Protein changes in tobacco pollen culture; a newly synthesized protein related to pollen tube growth. *J. Plant Physiol.* **130**, 307–314 (1987).

211. Metcalf, J. S., Barakate, A. & Codd, G. A. Inhibition of plant protein synthesis by the cyanobacterial hepatotoxin, cylindrospermopsin. *FEMS Microbiol. Lett.* **235**, 125–129 (2004).

212. Shemesh, T. *et al.* A model for the generation and interconversion of ER morphologies. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E5243-51 (2014).

213. Tolley, N. *et al.* Transmembrane domain length is responsible for the ability of a plant reticulon to shape endoplasmic reticulum tubules *in vivo*. *Plant J.* **64**, 411–418 (2010).

214. Wang, N. *et al.* Mechanism of membrane-curvature generation by ER-tubule shaping proteins. *Nat. Commun.* **12**, 568 (2021).

215. Wang, N. & Rapoport, T. A. Reconstituting the reticular ER network - mechanistic implications and open questions. *J. Cell Sci.* **132**, (2019).

216. Breeze, E. *et al.* A C-terminal amphipathic helix is necessary for the *in vivo* tubule-shaping function of a plant reticulon. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 10902–10907 (2016).

217. Lee, H. *et al.* An *Arabidopsis* reticulon and the atlastin homologue RHD3-like2 act together in shaping the tubular endoplasmic reticulum. *New Phytol.* **197**, 481–489 (2013).

218. Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).

219. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

220. Zaccai, N. R. *et al.* A *de novo* peptide hexamer with a mutable channel. *Nat. Chem. Biol.* **7**, 935–941 (2011).

221. Watson, R. T. & Pessin, J. E. Transmembrane domain length determines intracellular membrane compartment localization of syntaxins 3, 4, and 5. *Am. J. Physiol. Cell Physiol.* **281**, C215-23 (2001).

222. Lankford, S. P., Cosson, P., Bonifacino, J. S. & Klausner, R. D. Transmembrane domain length affects charge-mediated retention and degradation of proteins within the endoplasmic reticulum. *J. Biol. Chem.* **268**, 4814–4820 (1993).

223. Hain, T. M. *et al.* SPIRE-a software tool for bicontinuous phase recognition: application for plastid cubic membranes. *Plant Physiol.* **188**, 81–96 (2022).

224. Rosado, A. & Bayer, E. M. Geometry and cellular function of organelle membrane interfaces. *Plant Physiol.* **185**, 650–662 (2021).

225. Scorrano, L. *et al.* Coming together to define membrane contact sites. *Nat. Commun.* **10**, 1287 (2019).

226. Wu, Y. *et al.* Contacts between the endoplasmic reticulum and other membranes in neurons. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E4859–E4867 (2017).

227. Csordás, G. *et al.* Imaging interorganelle contacts and local calcium dynamics at the ER-mitochondrial interface. *Mol. Cell* **39**, 121–132 (2010).

228. Strasser, R. *et al.* Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol. J.* **6**, 392–402 (2008).

229. Pichler, H. & Emmerstorfer-Augustin, A. Modification of membrane lipid compositions in single-celled organisms - From basics to applications. *Methods* **147**, 50–65 (2018).

230. Guerfal, M. *et al.* Enhanced membrane protein expression by engineering increased intracellular membrane production. *Microb. Cell Fact.* **12**, 122 (2013).

231. Schoene, C., Fierer, J. O., Bennett, S. P. & Howarth, M. SpyTag/SpyCatcher cyclization confers resilience to boiling on a mesophilic enzyme. *Angew. Chem. Int. Ed Engl.* **53**, 6101–6104 (2014).

232. Kriechbaumer, V. ER microsome preparation in *Arabidopsis thaliana*. *Methods Mol. Biol.* **1691**, 117–123 (2018).

233. Araùjo, M., Hube, L. A. & Stasyk, T. Isolation of endocitic organelles by density gradient centrifugation. in *2D PAGE: Sample Preparation and Fractionation* (ed. Posch, A.) vol. 424 317–331 (Humana Press, 2008).

234. Weber, P. J. A., Weber, G. & Eckerskorn, C. Isolation of organelles and prefractionation of protein extracts using free-flow electrophoresis. *Curr. Protoc. Protein Sci.* **32**, 22.5.1-22.5.21 (2003).

235. Simon, K. S., Pollock, N. L. & Lee, S. C. Membrane protein nanoparticles: the shape of things to come. *Biochem. Soc. Trans.* **46**, 1495–1504 (2018).

236. Hering, J. *et al.* The rapid “teabag” method for high-end purification of membrane proteins. *Sci. Rep.* **10**, 16167 (2020).

237. Subbarao, G. V. & van den Berg, B. Crystal structure of the monomeric porin OmpG. *J. Mol. Biol.* **360**, 750–759 (2006).

238. Clarke, J. *et al.* Continuous base identification for single-molecule nanopore DNA sequencing. *Nat. Nanotechnol.* **4**, 265–270 (2009).

239. Hausjell, J., Halbwirth, H. & Spadiut, O. Recombinant production of eukaryotic cytochrome P450s in microbial cell factories. *Biosci. Rep.* **38**, BSR20171290 (2018).

240. Thoring, L., Dondapati, S. K., Stech, M., Wüstenhagen, D. A. & Kubick, S. High-yield production of “difficult-to-express” proteins in a continuous exchange cell-free system based on CHO cell lysates. *Sci. Rep.* **7**, 11710 (2017).

241. Idan, O. & Hess, H. Origins of activity enhancement in enzyme cascades on scaffolds. *ACS Nano* **7**, 8658–8665 (2013).

242. Betti, M. *et al.* Manipulating photorespiration to increase plant productivity: recent advances and perspectives for crop improvement. *J. Exp. Bot.* **67**, 2977–2988 (2016).

243. Meng, X. *et al.* Metabolic engineering *Saccharomyces cerevisiae* for de novo production of the sesquiterpenoid (+)-nootkatone. *Microb. Cell Fact.* **19**, 21 (2020).

244. Polturak, G. *et al.* Engineered gray mold resistance, antioxidant capacity, and pigmentation in betalain-producing crops and ornamentals. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 9062–9067 (2017).

245. Polturak, G. & Aharoni, A. "*La vie en rose*": Biosynthesis, sources, and applications of betalain pigments. *Mol. Plant* **11**, 7–22 (2018).

246. Duan, L. *et al.* Biosynthesis and engineering of kaempferol in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **16**, 165 (2017).

247. Arendt, P. *et al.* An endoplasmic reticulum-engineered yeast platform for overproduction of triterpenoids. *Metab. Eng.* **40**, 165–175 (2017).

248. Kim, J.-E. *et al.* Tailoring the *Saccharomyces cerevisiae* endoplasmic reticulum for functional assembly of terpene synthesis pathway. *Metab. Eng.* **56**, 50–59 (2019).

249. Jaramillo-Madrid, A. C., Lacchini, E. & Goossens, A. Within and beyond organelle engineering: strategies for increased terpene production in yeasts and plants. *Current Opinion in Green and Sustainable Chemistry* **33**, 100572 (2022).

250. Zhang, B. *et al.* Coordinated protein co-expression in plants by harnessing the synergy between an intein and a viral 2A peptide. *Plant Biotechnol. J.* **15**, 718–728 (2017).

251. Fu, J., Liu, M., Liu, Y., Woodbury, N. W. & Yan, H. Interenzyme substrate diffusion for an enzyme cascade organized on spatially addressable DNA nanostructures. *J. Am. Chem. Soc.* **134**, 5516–5519 (2012).

252. Veggiani, G. *et al.* Programmable polyproteams built using twin peptide superglues. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 1202–1207 (2016).

253. Gelvin, S. B. Integration of *Agrobacterium* T-DNA into the plant genome. *Annu. Rev. Genet.* **51**, 195–217 (2017).

254. Roux, A. *et al.* Membrane curvature controls dynamin polymerization. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 4141–4146 (2010).

255. Chen, L.-Q. *et al.* Sucrose efflux mediated by SWEET proteins as a key step for phloem transport. *Science* **335**, 207–211 (2012).

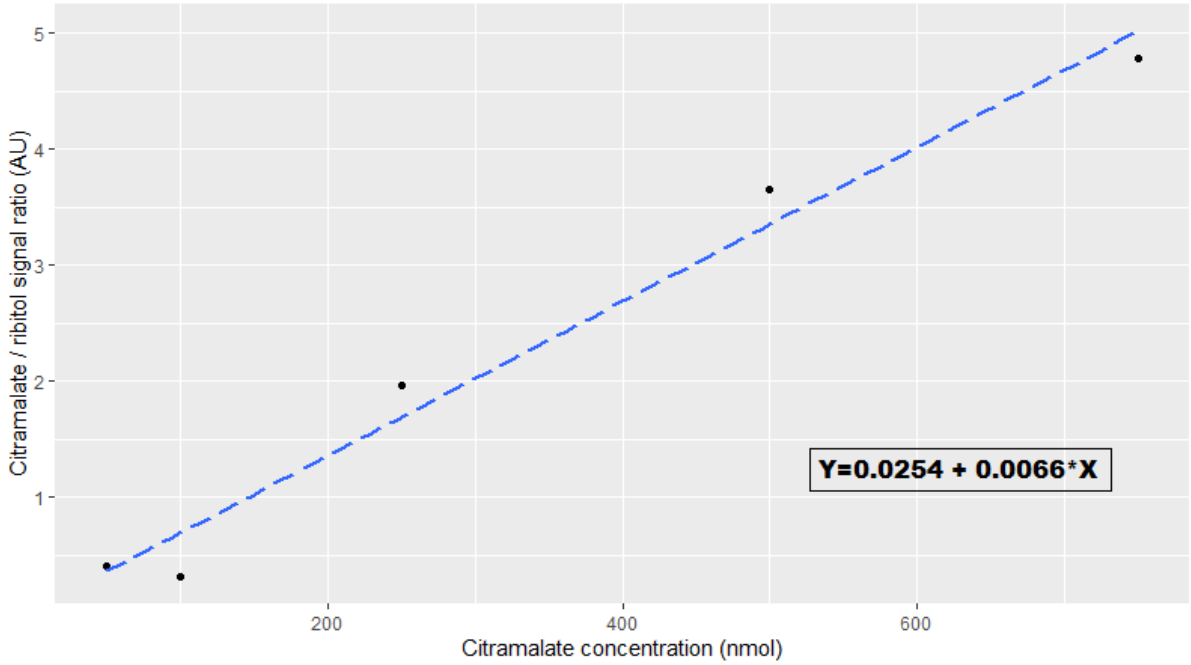
256. Ma, J. K.-C. *et al.* Molecular farming for new drugs and vaccines. Current perspectives on the production of pharmaceuticals in transgenic plants. *EMBO Rep.* **6**, 593–599 (2005).

257. Elbehri, A. Biopharming and the food system: Examining the potential benefits and risks. *AgBioForum* 18–25 (2005).

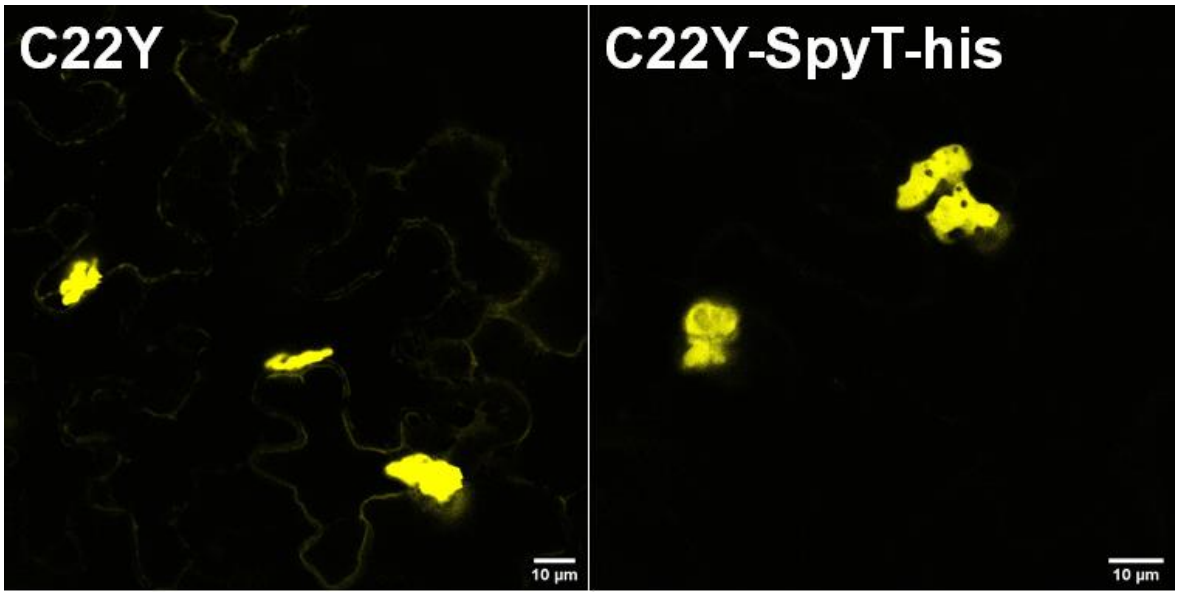
258. Sadre, R. *et al.* Cytosolic lipid droplets as engineered organelles for production and accumulation of terpenoid biomaterials in leaves. *Nat. Commun.* **10**, 853 (2019).

259. Liu, G.-S. *et al.* The yeast peroxisome: A dynamic storage depot and subcellular factory for squalene overproduction. *Metab. Eng.* **57**, 151–161 (2020).

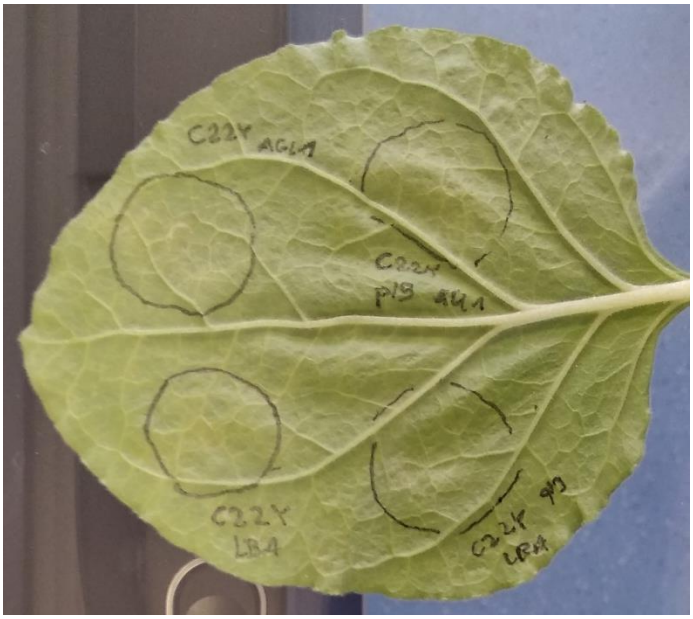
Appendix A: Supporting Information



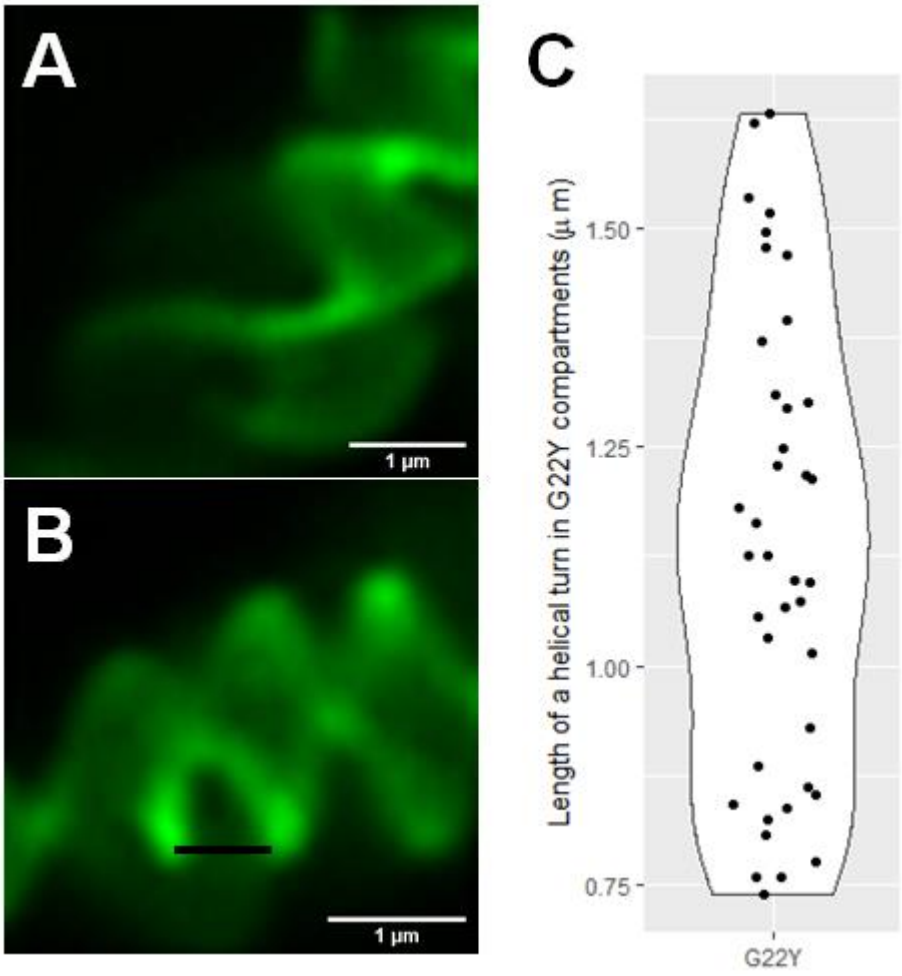
Supplementary Figure S1: The citramalate – ribitol standard curve. This was generated using a pure citramalate standard to aid the absolute quantification of citramalate in the samples shown in Chapter 5.2.5 and Figure 5.9. The linear regression of the values is shown as a blue dashed line with the respective linear equation in the black box.



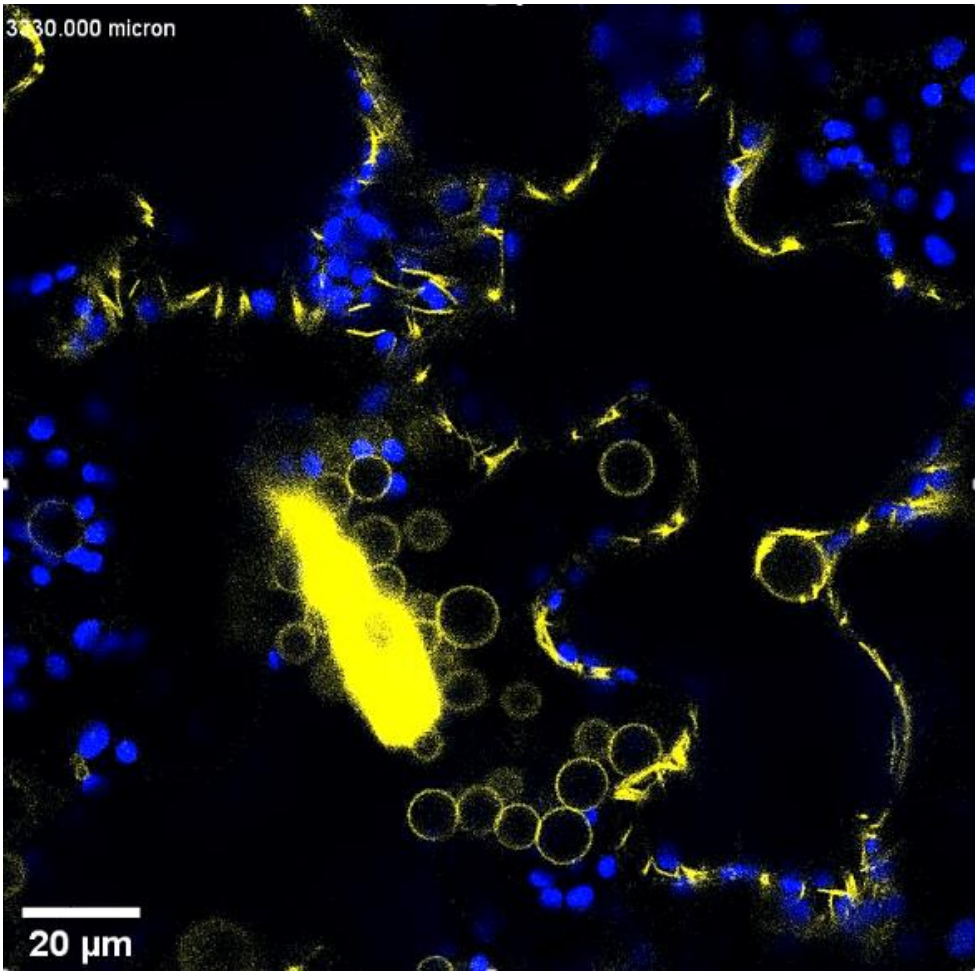
Supplementary Figure S2: Stable 2-week-old compartments. C22Y and C22Y-SpyT-his compartments remained stable for several weeks after transient transformation of mature tobacco leaves. Scale bars: 10 µm.



Supplementary Figure S3: Transiently transformed leaves show no negative macroscopic phenotypes. Image was taken 14 days after transient transformation with C22Y. Circled areas show infiltrated tissue.



Supplementary Figure S4: Helical structures observed in G22Y compartments. A&B: High-resolution airy-scanning confocal microscopy images of mature tobacco leaf epidermis captured 5 days after *A. tumefaciens*-mediated transformation. GFP signal is shown in green. Black bar shows length of a helical turn as measured. C: Measured values of helical turn lengths for G22Y compartments. n=39 Scale bars: 1 µm.



Supplementary Video SV1: Z-stack of 22Y structures. Video file is attached separately as part of Additional Materials. The above image is a representative snapshot of the Z-stack for use in absence of the video. 22Y (yellow) show a variety of structures, including the large compartments and large spherical swollen ER bodies. Blue: chloroplasts. Scale bar: 20 µm.

Parameter type	Parameter name	ANOVA p-value
Tubule edges	Length	0.673076412
Tubule edges	Volume	1
Tubule edges	Resistance	1
Tubule edges	Tortuosity	1
Tubule edges	Width_center	1
Tubule edges	Speed_local	1
Tubule edges	Speed_max	1
Tubule edges	Speed_global	1
Tubule edges	Flow_coherence	1
Tubule edges	Persistency	0.830678296
Tubule nodes	node_Omin_Omaj	1
Tubule nodes	node_Omid_Omaj	1
Tubule nodes	node_Omin_Omid	1
Tubule nodes	node_Speed	1
Tubule nodes	node_Persistency	0.425715629
Cisternae	cisternal_Node_degree	1
Cisternae	cisternal_Contrast_all	0.106037489
Cisternae	cisternal_Correlation_all	0.094236654
Cisternae	cisternal_Energy_all	1
Cisternae	cisternal_Homogeneity_all	0.095700979
Cisternae	cisternal_Speed_local	1
Cisternae	cisternal_Speed_global	1
Cisternae	cisternal_Flow_coherence	0.094521322
Cisternae	cisternal_MaxPersistency	1
Cisternae	cisternal_MeanPersistency	1
Cisternae	cisternal_MaxDistance	1
Cisternae	cisternal_MeanDistance	1
Cisternae	cisternal_Area	1
Cisternae	cisternal_Solidity	1
Cisternae	cisternal_Elongation	1
Cisternae	cisternal_Roughness	1
Cisternae	cisternal_circularity	1
Polygons	polygon_Area	0.002267661 (**)
Polygons	polygon_Solidity	1
Polygons	polygon_Circularity	1
Polygons	polygon_Roughness	1
Polygons	polygon_Elongation	0.000467863 (***)
Polygons	polygon_MaxDistance	0.005984266 (**)
Polygons	polygon_MeanDistance	0.006682543 (**)
Graph	Geff	1
Graph	Alpha	1

Supplementary Table S1: Individual ANOVA of every parameter used in AnalyzER testing in Chapter 3.2.5. Only four parameters have statistically significant ANOVAs (marked with asterisks) but when further analysed using Tukey's HSD test, neither had statistically significant difference between the WT and either of the groups, only between different compartment types. Due to the large number of tests performed, Bonferroni correction for multiple testing was used to reduce the multiple comparisons problem.