

EXPRESSION OF PARTICULATE METHANE
MONOOXYGENASE (pMMO) PROTEINS IN PLANTS
FOR METHANE DETOXIFICATION

Tatiana Spatola Rossi

St Hilda's College
Department of Plant Sciences
University of Oxford



Thesis submitted for the degree of
Doctor of Philosophy

Michaelmas Term, 2022

Supervisors: Dr Mark Fricker and Dr Verena Kriechbaumer

Abstract

Methane is a potent greenhouse gas which has contributed to approximately a fifth of global warming since pre-industrial times and its concentration continues to rise yearly in parallel to that of CO₂. Anthropogenic activities contribute to 70% of global emissions, with the main sources being the extraction and transport of fossil fuels, farming of livestock, waste management, rice agriculture, and biomass burning. Rice cultivation is a major source of anthropogenic emissions due to the flooded conditions of the soil and presence of organic matter which pose favourable conditions for biotic methanogenesis. In turn, the rice plants transport the greenhouse gas into the atmosphere through the porous aerenchyma tissue in a process termed “the chimney effect”. Strategies for mitigating rice emissions include modifying the watering regimes and the supplementation of organic additives. However, these practices are not always feasible for farmers and may introduce drawbacks such as lower grain yields and the emission of N₂O. In this thesis, the bioengineering of rice for methane mitigation is postulated. In nature, methanotroph bacteria use methane as their source of carbon and energy via methane monooxygenase enzymes (MMOs), and predominantly through particulate methane monooxygenase (pMMO). pMMO is encoded in an operon and forms a membrane-bound nonamer complex composed of three different protein subunits. The heterologous expression of pMMO has been deemed challenging due to its membrane localization and copper content. In this work the recombinant expression of pMMO in tobacco and Arabidopsis is presented as a proof of concept, with the further aim of bioengineering rice to contain pMMO and reduce emissions from paddy fields. The pmoC, pmoA and pmoB subunits were targeted to the endoplasmic reticulum (ER) individually, and gene stacking strategies were explored. The use of 2A self-cleaving peptides resulted in a successful gene stacking methodology to express all the subunits in the ER. The 2A pMMO polypeptide expression was greatly enhanced by co-expression with P19, and did not induce major disruption or modification to ER structure and dynamics. This thesis shows, for the first time, that the pMMO proteins can be co-

expressed in plant cells without apparent detrimental effects, and strategies for obtaining a functional enzyme in the recombinant system are investigated and discussed.

Acknowledgements

I would like to thank my supervisors, Dr Mark Fricker and Dr Verena Kriechbaumer, for being fundamental in the development and completion of this thesis. I am thankful to Verena for devising the project and for the guidance and knowledge with practical work. I would like to thank Mark for the help with his numerous complex software and helping me to develop skills in image analysis. I am also grateful to both for the time offered to discuss and analyze my results, and the many suggestions during writing, which were essential for me to shape the work in the form of a thesis and to grow as a scientist.

I would also like to thank Frances Tolmie for guiding me in the lab at the beginning, for help doing chloroplast transformation, and for being a very bright presence in the lab. I would like to acknowledge Charlotte Pain for help with lab related things, with microscopy and for the FRET-FLIM work, and the rest of my lab colleagues, especially Stefan Wojcik, Bisa Andov and Alastair McGinness for companionship during this work. I would also like to thank Tom Smith for the time dedicated to discussions regarding pMMO enzymatic assays with plant samples, and our initial attempts at testing activity in Sheffield.

My friendships created during the PhD really helped to shape this experience and make it memorable, especially Sofia, Claire, Henry, Razia, Tika and Gabriele, and also my friends back at home Mica, Dani, Yao and Herno.

I recognize the BBSRC, the Gen Foundation, and the Sir Richard Stapley Educational Trust for financial assistance.

Finally and very importantly I would like to thank my family for being extremely supportive throughout, without them I would not have been able to do this (Mamá, Papá, Carla, Abu, Nono, Ale, Martin, Lore, Franca y Matea).

Statement of authorship

The work presented in this thesis was carried out entirely by the author (Tatiana Spatola Rossi), with the exception of the generation of marker constructs already present in the Kriechbaumer lab (RFP-HDEL, GFP-HDEL, GFP-CNX, ST-RFP, RFP-ATPase, roGFP2 and roGFP2-HDEL – see table 2.3 for associated publications), the generation of Clover-pmoA and mCherry-pmoC, previously cloned by Verena Kriechbaumer from *Methylosinus trichosporium* OB3b DNA, and the acquisition of FRET-FLIM data which was carried out by Charlotte Pain and analysis of the raw data (generation of average lifetime values) by Verena Kriechbaumer.

Abbreviations

ANOVA	Analysis of variance
cDNA	Complementary deoxyribonucleic acid
CFP	Cyan fluorescent protein
CLSM	Confocal laser scanning microscopy
CNX	Calnexin
cv	Cultivar
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
FRET-FLIM	Förster resonance energy transfer-Fluorescence lifetime imaging
GFP	Green fluorescent protein
GM	Genetically modified
ICMs	Intracytoplasmic membranes
IPTG	Isopropyl β -D-1-thiogalactopyranoside
IRES	Internal ribosome entry site
LB	Luria-Bertani
LB/RB	Left border/ right border
MANOVA	Multivariate analysis of variance
MDH	Methanol dehydrogenase
MES	2-(N-Morpholino)ethanesulfonic acid
MMOs	Methane monooxygenases
mRNA	Messenger ribonucleic acid
MS	Murashige and Skoog medium
NADH	Reduced nicotinamide adenine dinucleotide
NADPH	Reduced nicotinamide adenine dinucleotide phosphate

NDH-2	Type II NADH:quinone oxidoreductase
NEB	New England biolabs
NIAB	National Institute of Agricultural Botany
OD	Optical density
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
pMMO	Particulate methane monooxygenase
PTGS	Post-transcriptional gene silencing
PSF	Point spread function
Q	Ubiquinone
QH₂	Ubiquinol
RFP	Red fluorescent protein
RNA	Ribonucleic acid
ROI	Region of interest
RT-PCR	Reverse transcription polymerase chain reaction
RTN1	Reticulon 1
RuMP	Ribulose monophosphate pathway
s.e.m	Standard error of the mean
SD	Standard deviation
sMMO	Soluble methane monooxygenase
SRP	Signal recognition particle
ST	Sialyltransferase
T-DNA	Transfer DNA
TAE	Tris Acetate-EDTA buffer
TGS	Transcriptional gene silencing
TM	Transmembrane domain
WT	Wild type

Table of Contents

1	Introduction	14
1.1	Overview	14
1.2	Methane emissions in the climate crisis	15
1.3	Methane in rice and wetland ecosystems	16
1.3.1	Methanogenesis.....	16
1.3.2	Methanotrophy	17
1.3.3	Plants act as a conduit for methane transport.....	19
1.4	Existing strategies towards methane emission mitigation in paddy fields.....	21
1.4.1	Agricultural management practices	21
1.4.2	Genetic modification of rice	22
1.5	Nature's methane catalysts: methane monooxygenases (MMOs)	23
1.5.1	Soluble methane monooxygenase (sMMO).....	23
1.5.2	Particulate methane monooxygenase (pMMO)	26
1.5.3	pMMO active site and physiology	28
1.5.4	pMMO kinetics	30
1.6	MMOs in biotechnology	32
1.6.1	Methane bioremediation	32
1.6.2	Recombinant expression of pMMO.....	33
1.7	Objectives and approach	35
2	Materials and methods.....	37
2.1	Plant material and growth conditions.....	37
2.2	Cloning.....	38
2.2.1	Gateway cloning	38
2.2.2	Cloning of single subunits with fluorescent tags	39

2.2.3	Construction of 2in1 vectors	43
2.2.4	Cloning of polypeptide vectors	46
2.3	Transformation of competent Agrobacterium cells	48
2.4	Transient transformation of tobacco	49
2.5	Generation of stable Arabidopsis plants.....	50
2.6	Confocal microscopy.....	51
2.7	Quantifying fluorescence intensity levels	52
2.7.1	Experimental design.....	52
2.7.2	Image processing	53
2.8	Redox ratio analysis	57
2.9	AnalyzER	59
2.10	FRET-FLIM.....	60
2.11	RT-PCR	61
2.11.1	RT-PCR of transiently expressed RFP-polypeptide in tobacco.....	61
2.11.2	RT-PCR of rice plants.....	62
2.12	Substrates for propylene epoxidation assays	63
2.12.1	Preparation of plant lysates	63
2.12.2	Reduction of duroquinone.....	63
3	Expression of pMMO proteins in the plant endoplasmic reticulum (ER)	64
3.1	Introduction	64
3.1.1	The role of the ER in protein synthesis and folding	64
3.1.2	The plant ER as a compartment for foreign protein production	65
3.1.3	Methods for testing protein-protein interactions.....	67
3.1.4	Methods for determination of membrane topology	69
3.1.5	Objectives	70
3.2	Results	72

3.2.1	Subcellular localization of pMMO proteins in tobacco cells.....	72
3.2.1.1	Protein structure and prediction of localization in Eukarya.....	72
3.2.1.2	PmoC and pmoA localize to the plant ER.....	77
3.2.1.3	PmoB localizes to the cytosol and to bright puncta in a signal peptide-dependent manner	82
3.2.1.4	PmoB puncta form ring structures which co-localize with mitochondria.....	86
3.2.2	PmoB does not relocate to the ER when co-expressed with pmoA.....	88
3.2.3	Targeting pmoB to the ER	90
3.2.3.1	Targeting pmoB via a plant signal peptide is successful but results in some protein escape to the vacuole	90
3.2.3.2	Addition of the HDEL retention signal produces accumulation of spBtrunc-RFP in the ER	94
3.2.3.3	The HDEL retention signal may marginally increase fluorescence of spBtrunc-GFP	96
3.2.3.4	Suppression of post-transcriptional gene silencing increases expression levels	99
3.2.4	Testing protein-protein interactions.....	100
3.2.4.1	Co-expression of pMMO subunits on the ER	100
3.2.4.2	Determination of protein-protein interactions with FRET-FLIM.....	101
3.2.4.3	ER retention of vacuolar pmoB in the presence of pmoA	109
3.2.5	Topology of pMMO proteins on the ER membrane.....	110
3.2.6	Co-expression of all three subunits using three-color imaging	113
3.3	Discussion	116
3.3.1	Bacterial pmoC and pmoA localize recombinantly to the plant ER.....	116
3.3.2	Bacterial pmoB can be targeted to the plant ER.....	117
3.3.3	Recombinant pMMO proteins interact with each other.....	119
3.3.4	Topology of pmoA.....	122

4	Strategies for gene stacking	126
4.1	Introduction	126
4.1.1	Gene stacking with multiple promoters	126
4.1.2	Gene stacking with a single promoter: polypeptides	128
4.1.3	Objectives	130
4.2	Results	132
4.2.1	Combinatorial design of “2in1” vectors for protein co-expression	132
4.2.2	Single cassette 2in1 vectors	134
4.2.3	Co-expression of 2in1 vectors with p19	135
4.2.4	Design and transcription of a pMMO polypeptide using self-cleaving peptides 138	
4.2.5	Co-expression and localization of subunits using 2A self-cleaving peptides..	140
4.2.6	Quantifying expression levels of pMMO polypeptides	143
4.2.7	Protein-protein interactions in pMMO polypeptides	146
4.2.8	Ratiometric expression of pMMO subunits in polypeptides	149
4.3	Discussion	152
4.3.1	Evaluation of co-expression using 2in1 vectors	152
4.3.2	Co-expression and ER targeting with 2A self-cleaving peptides	153
4.3.3	Interactions and stoichiometry in 2A polypeptides	155
5	Impact of pMMO on ER structure and dynamics.....	157
5.1	Introduction	157
5.1.1	Structure and movement of the plant ER.....	157
5.1.2	AnalyzER: a tool for the quantitative analysis of ER structure and dynamics	159
5.2	Results	162
5.2.1	Evaluation of the impact of polypeptides on the structure of the ER	162
5.2.2	Quantitative analysis of ER structure using AnalyzER	166

5.2.2.1	Experiment design.....	166
5.2.2.2	Co-expression with GFP-HDEL and skeletonization	167
5.2.2.3	Cisternal structure, polygonal areas and tubule length	172
5.2.2.4	Tubule morphology analysis	176
5.2.3	Quantitative analysis of ER movement and persistency.....	178
5.2.4	Network comparisons using multiple parameters.....	181
5.3	Discussion	185
5.3.1	Network morphology and texture metrics	185
5.3.2	The polypeptide does not cause constriction in tubules.....	186
5.3.3	Impact of pMMO proteins on dynamics and on the multivariate characterization of the network	188
6	Towards assays for pMMO in plants with stable expression of pMMO	190
6.1	Introduction	190
6.1.1	Measurement of the enzymatic activity of pMMO.....	190
6.1.2	Objectives	192
6.2	Results	195
6.2.1	Stable Arabidopsis plants expressing pMMO.....	195
6.2.2	Analysis of the impact of protein additions in the polypeptide	202
6.2.3	Design and transient expression of a modified polypeptide for increased functionality	205
6.2.4	Design of propylene epoxidation assays and sample preparation	208
6.2.5	Transgenic rice containing the pMMO polypeptide	210
6.3	Discussion	218
6.3.1	Creation of stable Arabidopsis plants containing pMMO	218
6.3.2	Testing enzymatic activity of pMMO in plants	218
6.3.3	The generation of rice containing pMMO	220

7	General discussion	222
7.1	Expression of pmoC, pmoA and pmoB in the plant ER	222
7.2	Assembly of the pMMO complex in the plant ER.....	224
7.3	Self-cleaving peptides for gene stacking.....	227
7.4	Impact of the single subunits and the pMMO polypeptide on ER structure and dynamics	229
7.5	Towards creating pMMO rice	232
7.5.1	Physiological aspects: reducing equivalents, copper requirement and methanol production	232
7.5.2	Obstacles and successes in the development of GM rice	235
8	Future work.....	237
8.1	Testing enzymatic activity of RFP-poly and RFP-polyM <i>in vitro</i>	237
8.2	Performing enzyme assays <i>in vivo</i>	238
8.3	Expression of pMMO in the chloroplast	239
8.3.1	Advantages of chloroplast transformation	239
8.3.2	Localizing transgenes to the thylakoid membrane.....	240
8.3.3	Preliminary work on the chloroplast transformation with pMMO	241
9	Supplementary data	245
9.1	Chapter 2 – Materials and methods.....	245
9.2	Chapter 3 – Expression of pMMO proteins in the plant endoplasmic reticulum (ER) 249	
9.3	Chapter 5 – Impact of pMMO on ER structure and dynamics.....	256

1 Introduction

1.1 Overview

Methane gas is a significant contributor to global warming occurring as the product of anthropogenic activities since pre-industrial times (Nisbet *et al.*, 2014). Although it is present in much lower atmospheric concentrations than CO₂, it has a significantly higher global warming potential but a shorter lifetime (IPCC, 2014). These characteristics make methane an attractive target for tackling global warming in the short-term (IPCC, 2014; Saunio *et al.*, 2016a), especially since current emission scenarios predict global warming to exceed the 1.5°C limit in only a decade (IPCC, 2021). The main sources of anthropogenic methane emissions include the extraction and transport of fossil fuels, livestock farming, rice agriculture, waste management and biomass burning (Dlugokencky *et al.*, 2011; Jackson *et al.*, 2020). Rice agriculture contributes to around 9% of anthropogenic methane emissions (Saunio *et al.*, 2016b). Emission mitigation strategies in rice have been investigated, such as water management and breeding of cultivars with lower gas transport capacities (Minamikawa *et al.*, 2006; Hussain *et al.*, 2015; Bharali *et al.*, 2017; Bhattacharyya *et al.*, 2018). However, alterations to management practice is not always feasible for local farmers or in certain geographic conditions, such as reducing watering in rainfed and deepwater rice (Minamikawa *et al.*, 2006; Hussain *et al.*, 2015). Bioengineering rice for lower methane production could be key to combating emissions globally and efficiently. Only one example has been reported in the literature with genetically modified rice to produce a “high starch and low emissions” variety, which reduced emissions by 22-51% (Su *et al.*, 2015; Du *et al.*, 2021). Although this is a promising start, parallel engineering approaches, for example using recombinant enzymes that detoxify methane, could further drive emissions to lower levels. In this respect, methanotroph bacteria naturally break down methane using particulate methane monooxygenase (pMMO), a membrane-bound enzyme complex which is the predominant

methane catalyst found in nature (Murrell *et al.*, 2000). The heterologous expression of pMMO has resulted challenging in the past (Gou *et al.*, 2006; Lu *et al.*, 2019; Koo and Rosenzweig, 2021). However, expression of pMMO in plant cells has not been attempted so far. Nevertheless, the presence of membranous compartments with a large surface area such as the endoplasmic reticulum could result in a feasible bioengineering strategy for the recombinant expression of pMMO in plants, and creation of low emission rice varieties.

1.2 Methane emissions in the climate crisis

Greenhouse gas emissions as a product of industrialisation have increased in recent decades to unprecedented levels compared to at least the last 800,000 years (IPCC, 2014), causing an increase in the global average temperature so far of 1.1°C since pre-industrial levels (IPCC, 2021). The single largest contributor to global warming is carbon dioxide (CO₂), and great investment and re-structuring of the economy, primarily in the energy sector, is still required in order to avoid extreme climate change scenarios (IPCC, 2014). The United Nations proposed in the Paris Agreement in 2015 to keep to a target of a 1.5°C or well below 2°C temperature increase, in order to avoid catastrophic climate scenarios, by setting net-zero economy targets by 2050. However, in spite of substantial measures taken by some countries, trends indicate that carbon is still being emitted into the atmosphere at a rate which is not compatible with this target, instead it is estimated that the 1.5°C increase in global warming will be crossed in the early 2030s (IPCC, 2021), and some scenarios project a 3°C increase by the end of the century (Rogelj *et al.*, 2016), which would contribute to extreme damage to humans, animals, and ecosystems.

Methane (CH₄) is the second largest contributor to global warming, accounting for approximately a fifth of the increase in radiative forcing since pre-industrial times (Nisbet *et al.*, 2014). Concentrations have increased dramatically from around 700 ppb in the mid-18th century to 1,890 ppb in 2021 (Dlugokencky *et al.*, 2011; Dlugokencky, 2021). The rate of

emissions experienced slower growth between 2000 and 2007, but have resumed rapid growth since 2007 (Nisbet *et al.*, 2014). Methane has a shorter atmospheric lifetime of 12.4 years compared to CO₂ (with a lifetime of hundreds of years), but a greater global warming potential, by a factor of 84 compared to CO₂ over a 20-year time frame (IPCC, 2014). Despite its lower atmospheric concentration, it is thus a very potent greenhouse gas, but arguably represents a good target for short-term climate change action (IPCC, 2014; Saunio *et al.*, 2016a). There are numerous sources of CH₄ emissions which include natural wetlands, enteric fermentation, paddy fields, thawing permafrost, gas, coal and oil industries, landfills, and burning biomass, amongst others (Nisbet *et al.*, 2014). Taken together, two-thirds of CH₄ emissions arise from anthropogenic activities, and in great measure from the agriculture and energy sectors (Saunio *et al.*, 2016a). Anthropogenic methane emissions therefore constitute a plausible focus for intervention through the modification of human practices and the development of sustainable solutions, alongside changes implemented to curb CO₂ emissions. The mitigation of human derived methane emissions has been largely unexplored compared to the mitigation of CO₂ and presents an opportunity to significantly impact climate change in the short-term, to avoid extreme climate scenarios.

1.3 Methane in rice and wetland ecosystems

1.3.1 Methanogenesis

Natural wetlands and rice agriculture are two of the main contributors towards global methane emissions. Natural wetlands are the single largest contributor to the total methane budget, producing around 30% of total global emissions (Nisbet *et al.*, 2014; Saunio *et al.*, 2016b; Poulter *et al.*, 2017). Rice cultivation contributes significantly to emissions, with some studies reporting a contribution of 10% to total emissions (Bhattacharyya *et al.*, 2018), whilst others report 2.5% of total emissions (Kritee *et al.*, 2018), or 9% of total anthropogenic emissions (Saunio *et al.*, 2016b). The difficulty in assigning a specific percentage of contribution lies on

the regional and temporal variability in contribution to emissions from the different sources, as well as the different methods to measure emissions and extrapolate to calculate a global average.

Methane in rice and wetland ecosystems is produced as a result of waterlogged soils and the presence of organic matter (Serrano-Silva *et al.*, 2014). These conditions favour the process of biotic methanogenesis, which takes place in anaerobic conditions through four successive stages (Le Mer and Roger, 2001; Chandra *et al.*, 2012). Firstly, in the stage of hydrolysis, bacteria break down complex organic molecules (carbohydrates, proteins and lipids) into smaller organic molecules and monomers (i.e. sugars, amino acids, and peptides). Secondly, during acidogenesis, fermentative bacteria convert the smaller organic molecules into volatile fatty acids (such as propionate and butyrate), alcohols, H₂ and CO₂. Subsequently, the metabolites from the previous stage are transformed into H₂, CO₂ and acetate in the acetogenesis stage. Finally, methanogenesis is carried out by methanogenic archaea (methanogens) where either H₂ and CO₂ or acetate are converted into CH₄ during hydrogenotrophic or acetogenic methanogenesis, respectively (Le Mer and Roger, 2001; Chandra *et al.*, 2012). Methanogens are strict anaerobes and require low redox potentials to grow, therefore the flooded conditions of the soil which create an anoxic environment pose favourable conditions for their growth and promote high levels of methanogenesis (Serrano-Silva *et al.*, 2014).

1.3.2 Methanotrophy

Closely coupled to the production of methane, wetlands and paddy fields also harbour a group of bacteria, termed methanotroph bacteria or methanotrophs, which consume methane using it as their sole source of carbon and energy, in a process called methanotrophy (Serrano-Silva *et al.*, 2014). Methanotrophs are aerobic Gram-negative bacteria which inhabit the interface between the anoxic and oxic layers of methane rich soils. In the first step of their metabolism, methane is oxidised to methanol, a reaction which is catalysed by the enzyme methane

monooxygenase (MMO) (Lieberman and Rosenzweig, 2004). This enzyme has two distinct forms: a cytosolic form termed soluble methane monooxygenase (sMMO), and a membrane bound form termed particulate methane monooxygenase (pMMO) (Lieberman and Rosenzweig, 2004) (Figure 1.1). The majority of methanotrophs express only pMMO, whereas very few species contain only sMMO, and a subgroup contains the genes for both pMMO and sMMO (Sirajuddin and Rosenzweig, 2015). In the latter case, the expression of one or other enzyme depends on the copper levels in the environment: sMMO is expressed only when copper levels are low (below 1 μM) and pMMO is expressed when copper levels are higher (Sirajuddin and Rosenzweig, 2015). Following the reaction by MMO, methanol is converted to formaldehyde by methanol dehydrogenase, and formaldehyde is further oxidised to CO_2 , releasing energy, or assimilated through the ribulose monophosphate cycle or the serine pathway (Semrau *et al.*, 2018).

It is estimated that approximately 80% of the methane produced in the soil is consumed by methanotrophs (Serrano-Silva *et al.*, 2014), however, the net release of the greenhouse gas is still large enough for wetlands and rice fields to constitute major sources of global emissions.

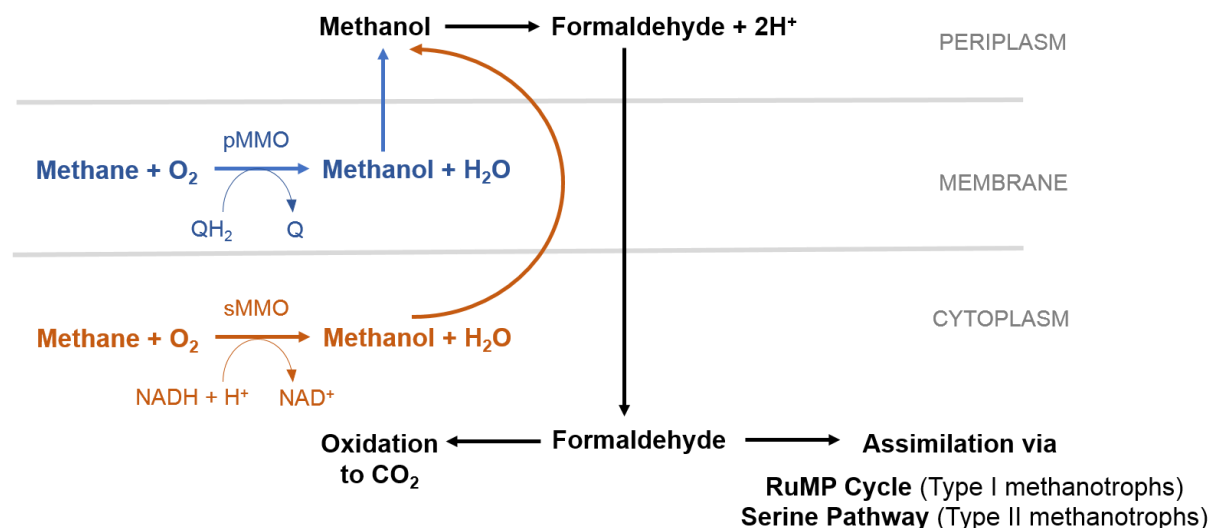


Figure 1.1. Methanotroph metabolism.

Pathways for the oxidation of methane by pMMO (blue) and sMMO (orange). Methane is converted to methanol by pMMO present in the membrane, using reducing equivalents most likely deriving from ubiquinol (QH₂), or by cytosolic sMMO using NADH as a reducing agent.

Methanol is then converted to formaldehyde by methanol dehydrogenase in the periplasmic space, and formaldehyde is further oxidized to CO₂, releasing energy, or assimilated via the ribulose monophosphate cycle (RuMP) or the serine pathway for the production of biomass. Modified from (Kalyuzhnaya *et al.*, 2015).

1.3.3 Plants act as a conduit for methane transport

In both natural wetlands and in paddy fields plants play a significant role in the production of methane and in its transport from the soil into the atmosphere. Plants release organic matter into the soil via exudation from the roots, and these root exudates serve as a substrate for hydrolytic bacteria in the first stage of methanogenesis (Bhullar *et al.*, 2013). Additionally, wetland and rice plants develop a porous tissue termed aerenchyma, which serves to supply oxygen to the submerged tissues such as the roots. However, the aerenchyma also acts as a conduit for methane, which is transported from the roots where it is absorbed, upward through the plant until its release into the atmosphere, a process termed the “chimney effect” (Bhullar *et al.*, 2013) (Figure 1.2).

The mechanism of methane transport through rice plants has been studied in detail. Nouchi *et al.* (1990) observed that dissolved methane in the soil diffuses into the roots, is then gasified and transported through the aerenchyma, and released primarily through micropores in the leaf culm and leaf sheath, and secondarily through stomata in the leaf blade (Nouchi *et al.*, 1990). The concentration of methane inside the medullar cavity of field grown rice was found to be 2900 times that of ambient air (Nouchi *et al.*, 1990), highlighting the scale of the chimney effect taking place. The contribution of plant-mediated transport to total methane emissions in paddy fields is very significant: Butterbach *et al.* (1997) reported that 90% of released methane is transported through the plants, with different varieties presenting diverse gas transport capacities which were associated with varying aerenchyma morphology (Butterbach-Bahl *et al.*, 1997). The remaining gas is transported either through diffusion or ebullition. Plant species in natural wetlands also harbour varying amounts of methane transport capacity, and studies have reported that the proportion of plant-mediated transport is between 30% and 100% in these environments (Koelbener *et al.*, 2010; Bhullar *et al.*, 2013). These observations indicate

that plants play an important role in funnelling global methane emissions and are therefore a suitable point of action to apply mitigation strategies.

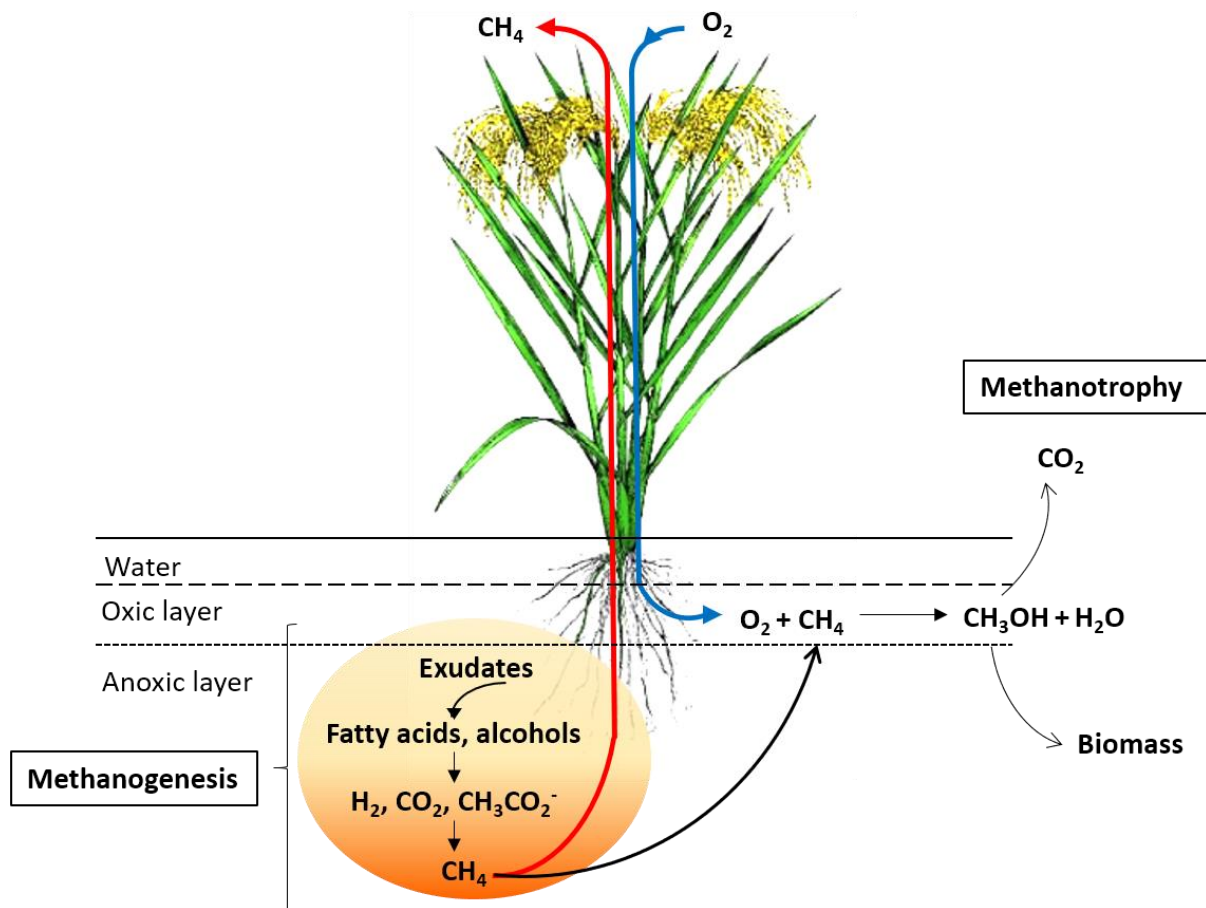


Figure 1.2. Methane production, consumption and transport in paddy fields.

Methane is produced in the anoxic layer of the soil by a consortium of fermentative and methanogenic microorganisms using root exudates as substrates. Methane is transported through the plant's aerenchyma into the atmosphere (termed the chimney effect, shown in red), whilst oxygen is transported through the aerenchyma in the opposite direction (shown in blue). In turn, oxygen and methane are used as substrates for methanotrophy in the oxic layer of the soil. Modified from (Philippot and Hallin, 2011).

1.4 Existing strategies towards methane emission mitigation in paddy fields

1.4.1 Agricultural management practices

Rice is a food staple for billions of people worldwide, and its cultivation constitutes a major source of methane emissions (Van Groenigen *et al.*, 2012). Greenhouse gas emissions from rice cultivation are expected to continue to grow as the demand for food increases each year with the increase in human population (Minamikawa *et al.*, 2006). Moreover, projections taking into account a warmer climate and increased CO₂ levels in the atmosphere predict that methane emissions of rice per kg will approximately double by 2100 (Van Groenigen *et al.*, 2012). Therefore, strategies to reduce emissions from paddy fields are essential for sustainable agriculture and to mitigate human induced climate change.

One of the most promising strategies to mitigate emissions from the crop is to modify the irrigation regime. Many studies exist regarding the impact of irrigation regimes on rice emissions. As an example, one study investigated the effects of three irrigation levels on emissions in long term field trials in China using two different rice varieties, one of them developed as a drought-resistant variety (Sun *et al.*, 2016). They found that lower amounts of irrigation decreased methane emissions in both varieties, however, grain yield was also reduced in the non-drought-resistant variety (Sun *et al.*, 2016). Reviews on this topic are very positive in stating that water management practise can effectively reduce that amount of emissions (such as mid-season drainage and alternate wetting and drying), and that further research and policies should be applied to implement these changes (Minamikawa *et al.*, 2006; Hussain *et al.*, 2015). However, drawbacks include potential trade-offs with grain yields, adapting new management practices to local farmers' capabilities, and the exclusion of rainfed or deepwater rice from these modifications (Minamikawa *et al.*, 2006; Hussain *et al.*, 2015). Moreover, some studies have found that reducing irrigation through alternate wetting and drying practices resulted in much higher levels of emissions of nitrous oxide (N₂O), another greenhouse gas which, even

if present at much lower concentrations than CH₄, has an even higher global warming potential and is long-lived, and would not be sustainable in the long-term (Kritee *et al.*, 2018).

Another management practice which has come under review is the application of biochar, a form of carbonised organic matter achieved through pyrolysis, to paddy fields. In several studies, the addition of rice-straw derived biochar was shown to significantly reduce methane emissions in paddy field soils (Han *et al.*, 2016). Biochar is thought to alter the physicochemical conditions of the soil favouring the growth of methanotrophs whilst inhibiting methanogenesis (Han *et al.*, 2016). This addition, however, may be costly for farmers who lack the suitable equipment to carry out pyrolysis in the necessary conditions (Kavitha *et al.*, 2018).

Other studies and policies aim to identify and promote the use of rice varieties which produce lower amounts of emissions. Some studies have shown a relationship between gas transport capacity and morphological and physiological traits such as aerenchyma size, amount of root oxidation, biomass production and emission in different growth stages, which provide useful information for potential selective breeding (Butterbach-Bahl *et al.*, 1997; Setyanto *et al.*, 2004; Bharali *et al.*, 2017; Bhattacharyya *et al.*, 2018).

1.4.2 Genetic modification of rice

Policies aimed at changing rice varieties or management practices are promising measures towards decreasing emissions in the short-term. However, the efficacy of each treatment is highly dependent on the geography of the region, the feasibility for farmers, and avoiding any reduction in crop yield. Genetic modification of rice to detoxify methane directly could provide a more homogeneous and more easily implemented solution.

One study produced transgenic rice containing “SUSIBA2”, a plant-specific transcription factor which modulates the expression of sugar-inducible genes (Su *et al.*, 2015). Transformants which overexpressed SUSIBA2 in the seeds and stems produced a higher allocation of photosynthates to the above ground biomass compared to roots, and resulted in increased biomass and starch in the seeds and stem, as well as in decreased CH₄ emissions,

potentially due to reduced root exudation (Su *et al.*, 2015). Recently, field trials were carried out to study the efficiency of SUSIBA2 “high-starch low-methane-emission” rice in decreasing emissions in different growing seasons and in two different varieties (*japonica* and *indica* rice) (Du *et al.*, 2021). It was found that SUSIBA2 *japonica* rice reduced emissions by around 51% in both the summer and autumn seasons, and by around 22% in the winter season, and SUSIBA2 *indica* rice reduced emissions by around 31% in the normal growing season (Du *et al.*, 2021), resulting in promising outcome.

Interestingly, a different study carried out a meta-analysis which concluded that larger root biomass produced a downregulation of methane fluxes (Chen *et al.*, 2021). Moreover, a field trial with a mutant of the OsRGA1 gene (*rga1*), which regulates photosynthate allocation to roots, showed that the mutant presented a lower belowground biomass but larger quantities of root exudates compared to the wild-type, and produced higher methane emissions and lower yield (Chen *et al.*, 2021). These findings show that a lower root biomass does not necessarily produce lower CH₄ emitting rice.

In spite of these pioneering efforts, the genetic modification of rice for methane mitigation remains largely unexplored, and paddy fields continue to be one of the main anthropogenic sources of methane emissions into the atmosphere. There is thus currently a need for further innovation and development of methane mitigation or methane capture technologies which could effectively tackle the issue.

1.5 Nature’s methane catalysts: methane monooxygenases (MMOs)

1.5.1 Soluble methane monooxygenase (sMMO)

Methane monooxygenases (MMOs) are nature’s only methane catalysts, and are able to break the very strong C-H bond ($\Delta H_{298} = 105$ kcal/mol) under ambient temperature and pressure conditions, a reaction otherwise carried out at very high temperatures and pressure using

chemical catalysts (Lieberman *et al.*, 2003; Sirajuddin and Rosenzweig, 2015). MMOs are expressed by methanotroph bacteria and have been of interest in research for decades for their potential applications in bioremediation, conversion of CH₄ to liquid fuels and chemicals, and for the development of biomimetic catalysts (Sirajuddin and Rosenzweig, 2015). Two types of MMOs are present in methanotrophs: soluble methane monooxygenase (sMMO) and particulate methane monooxygenase (pMMO), which both catalyse the conversion of methane to methanol with the following reaction (Koo and Rosenzweig, 2021),



sMMO is the least ubiquitous MMO found in nature as very few methanotrophs possess only sMMO, whereas a subset contain both sMMO and pMMO but only express sMMO when the copper:biomass ratio is very low (i.e. less than 1 µM copper in the growth medium) (Sirajuddin and Rosenzweig, 2015). However, given its soluble nature, the enzyme is easier to purify and is more stable after purification, therefore, its structure and reaction mechanism have already been elucidated (Lawton and Rosenzweig, 2016). sMMO is a cytosolic protein complex encoded in an operon containing 6 genes (*mmoXYBZDC*), where three of these genes code for the catalytic enzyme or hydroxylase component (*XYZ*), *mmoB* encodes a regulatory protein, *mmoC* codes for a reductase component, and *mmoD* also referred to as “*orfY*” is of unknown function (Murrell *et al.*, 2000). The enzyme also contains two associated genes: *mmoG*, encoding a putative GroEL chaperone, and *mmoR*, a transcriptional activator, which are necessary for sMMO expression in methanotroph cells (Scanlan *et al.*, 2009). sMMO is a metalloenzyme containing a di-iron centre in the alpha subunit of the hydroxylase. The hydroxylase is reduced by the reductase component (protein C), which in turn is reduced by NADH (Figure 1.3).

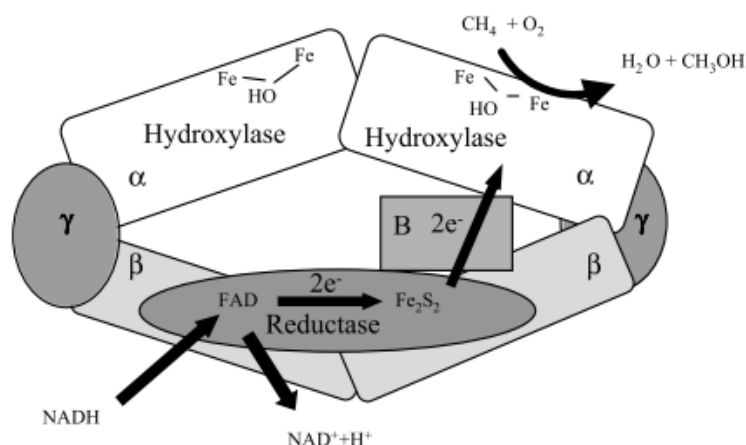


Figure 1.3. Model of sMMO structure.

Model of the sMMO complex containing a hydroxylase component (a dimer of the alpha, beta and gamma subunits) with a di-iron centre where conversion of methane into methanol takes place. The complex also contains protein B, and protein C (FAD-containing reductase) which receives electrons from NADH. Taken from (Murrell *et al.*, 2000).

The most commonly studied species of methanotrophs are *Methylococcus capsulatus* (Bath) and *Methylosinus trichosporium* OB3b. The turnover number of sMMO for methane oxidation in cells of *Methylosinus trichosporium* OB3b is at least 3.5 s^{-1} , and the turnover number of purified sMMO from *Methylosinus trichosporium* OB3b is 3.7 s^{-1} , whereas purified sMMO from *Methylococcus capsulatus* (Bath) has a lower turnover number of $0.2\text{-}1 \text{ s}^{-1}$ (reviewed in Sirajuddin and Rosenzweig, 2015). The enzyme thus shows relatively slow kinetics, compared for example with enzymes from central metabolism which have an average turnover numbers of 79 s^{-1} (Lawton and Rosenzweig, 2016). However, methanotrophs overcome the slow rates of conversion by highly expressing the MMO enzymes (Lawton and Rosenzweig, 2016). The Michaelis constant (K_m) for methane of sMMO from *Methylococcus capsulatus* (Bath) is $0.94 \mu\text{M}$ and for sMMO from *Methylosinus trichosporium* OB3b is $12 \mu\text{M}$ (Sirajuddin and Rosenzweig, 2015). In addition to metabolising methane, sMMO was found to have a wider substrate specificity than pMMO, oxidising C1-C8 alkanes, alkenes, aromatic compounds and halogenated hydrocarbons (Sirajuddin and Rosenzweig, 2015).

1.5.2 Particulate methane monooxygenase (pMMO)

Particulate methane monooxygenase (pMMO) is the predominant methane catalyst found in nature, present in almost all species of methanotrophs. pMMO is an integral membrane protein complex composed of three subunits which are encoded in an operon regulated by a single promoter: *pmoC*, *pmoA* and *pmoB* (Murrell *et al.*, 2000). The three subunits, with approximate molecular weights of 23 kDa, 26 kDa, and 45 kDa form a trimeric complex in a ratio 1:1:1 (Murrell *et al.*, 2000). Crystallography data shows that the quaternary structure of the complex consists of a nonamer, with the trimeric complex repeated three times (Lieberman and Rosenzweig, 2005) (Figure 1.4). The nonamer has also been observed in a physiologically relevant context using electron microscopy of active membranes containing pMMO (Kitmitto *et al.*, 2005). The crystal structures show that the *pmoC* and *pmoA* subunits are composed almost entirely of transmembrane domains, whereas *pmoB* possesses two transmembrane helices which separate two soluble domains that face the periplasm of methanotroph bacteria (Lieberman and Rosenzweig, 2005). Crystal structures from four different methanotroph genera have been solved (Table 1.1). The structure from the methanotroph species used in this work has the Protein Data Bank identification code (PDB ID) 3CHX.

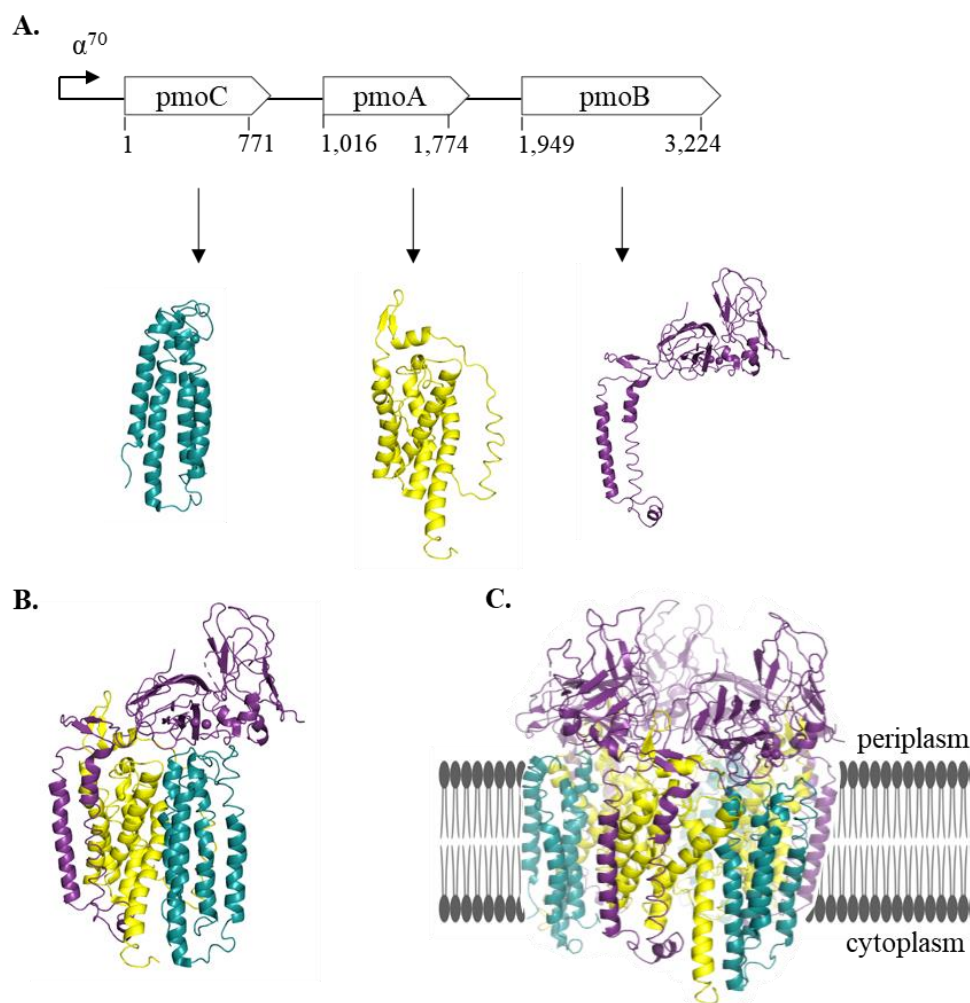


Figure 1.4. Structure of pMMO.

(A) Diagram of the pMMO operon containing pmoC, pmoA, and pmoB separated by two intergenic sequences (numbers indicate the base pairs from the operon of *Methylosinus trichosporium* OB3b (accession number: U31650.2)), and crystal structure of each protein subunit. (B) Crystal structure of the heterotrimer (PDB ID 3CHX), and the nonamer (C) inserted into the membrane, with the beta barrels from pmoB facing the periplasm.

Table 1.1. Crystal structures of the pMMO complex from different methanotroph genera.

Structures available for the pMMO complex showing the organism of origin, their identification code (PDB ID) and the date of publication.

Organism	PDB ID	Date
<i>Methylosinus trichosporium</i> OB3b	3CHX	2008
<i>Methylocystis</i> sp. strain M	3RFR	2011
<i>Methyломicrobium alcaliphilum</i> 20Z	6CXH	2018
<i>Methylococcus capsulatus</i> (Bath)	3RGB	2011

1.5.3 pMMO active site and physiology

The pMMO complex forms part of an electron transport chain in the methanotroph membranes. The enzyme utilises reducing power which is thought to derive indirectly from NADH, however, it is yet unclear what constitutes the direct source of electrons to the enzyme (Sirajuddin and Rosenzweig, 2015). A likely theory is that electrons are provided by ubiquinol produced by a type 2 NADH:quinone oxidoreductase (NDH-2) (Figure 1.5) (Sirajuddin and Rosenzweig, 2015). This is supported by *in vitro* experiments which show that NADH can act as reducing power and elicit activity in methanotroph whole-cells and in membranes containing pMMO, but not in samples containing solubilized pMMO (Shiemke *et al.*, 1995). Instead, duroquinol (a synthetic analogue of plastoquinol) can act as an effective reducing agent for solubilised pMMO (Shiemke *et al.*, 1995).

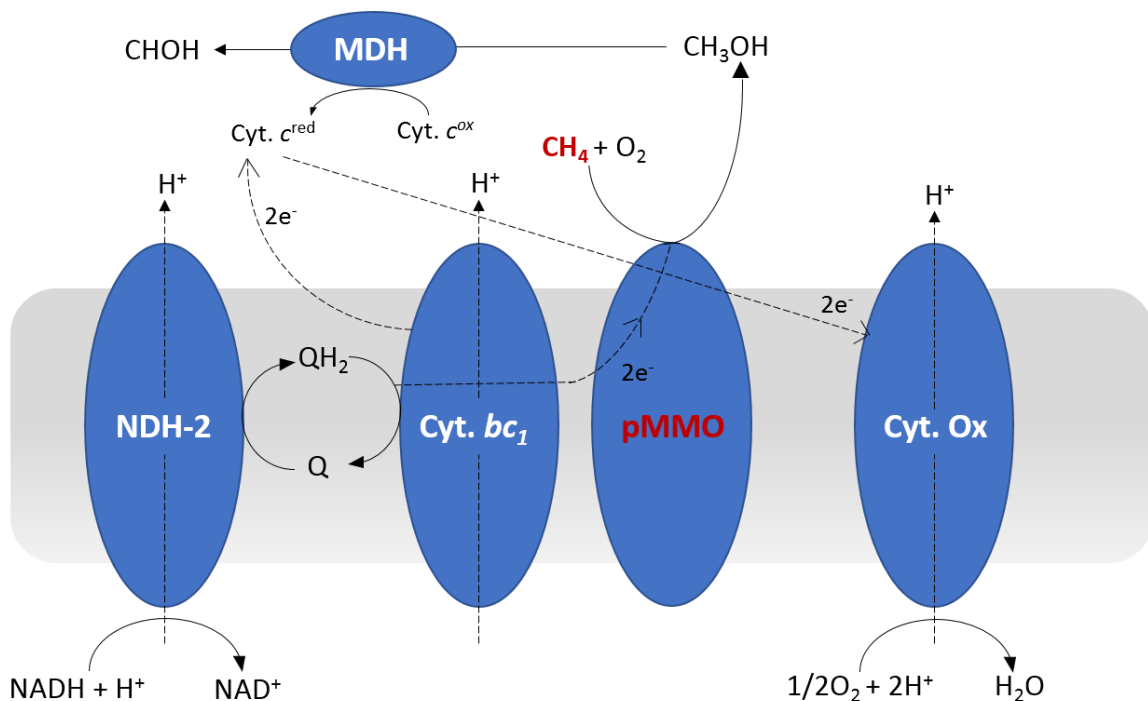


Figure 1.5. A proposed pathway of the respiratory chain in methanotrophs expressing pMMO.

The electron donor of pMMO is not yet known, however, some evidence suggests that electrons derive from ubiquinol (QH₂) recycled by a type 2 NADH dehydrogenase (NDH-2) and transferred via cytochrome *bc*₁ (Cyt. *bc*₁). The electron transport chain drives the translocation of protons to the periplasm which then allows ATP synthesis. Methanol produced by pMMO is oxidized to formaldehyde by methanol dehydrogenase (MDH) in the periplasm,

accompanied by the reduction of the water-soluble electron carrier cytochrome *c*. Cytochrome *c* then transfers electrons to cytochrome *c* oxidase (Cyt. Ox) which drives the reduction of O₂ to H₂O, transporting further protons into the periplasm. Modified from (Semrau *et al.*, 2010).

In contrast to sMMO, the difficulty in establishing some aspects of pMMO physiology is in part due to the instability of the enzyme when removed from the membrane and the significant loss of activity in the different stages of purification (Choi *et al.*, 2003; Sirajuddin and Rosenzweig, 2015). Further uncertainties remain regarding the nature of the active site and catalytic mechanism of the enzyme, due to the relatively low resolution of the crystal structures and the inability so far to express pMMO in *E. coli*, which could facilitate mutagenesis studies. pMMO is a copper dependent enzyme which contains 2 copper sites in most methanogen species: one site, denoted as Cu(II)_B, is located in the soluble domain of pmoB and is coordinated by three Histidine residues; the other site, Cu(II)_C, is located in the pmoC subunit and is coordinated by two Histidine and one Asparagine residues (Ross *et al.*, 2019). Initially, studies indicated that the active site was at Cu(II)_B, modelled as a dicopper centre, especially since an engineered soluble version of pmoB without its transmembrane domains appeared to show methane oxidation activity (Balasubramanian *et al.*, 2010). More recently, quantum refinement of the crystallography data has shown that pMMO possesses only monocopper sites (Cao *et al.*, 2018), and further experiments on the soluble pmoB form showed that it did not exhibit methane oxidation activity (Ross *et al.*, 2019). Instead, evidence is presented towards Cu(II)_C being the active site of catalysis (Ross *et al.*, 2019). Another study from the Rosenzweig group has employed nanodiscs and native top-down mass spectrometry that also supports the presence of monocopper sites, and a prominence of the Cu_C site during catalysis (Ro *et al.*, 2019). The nature and nuclearity of the copper sites is essential for the determination of a reaction mechanism, which is still unresolved (Ross and Rosenzweig, 2017). Recent computational studies have simulated potential reaction mechanisms based on activation of O₂ in the monocopper Cu(II)_C site (Morris, 2020; Peng *et al.*, 2021), and further spectroscopy studies have reaffirmed the presence of only monocopper sites and the Cu(II)_C as the likely site of methane activation (Cutsail *et al.*, 2021; Jodts *et al.*, 2021). However, a recent cryo-EM

structure of pMMO from *Methylococcus capsulatus* Bath suggests the presence of additional copper clusters and the presence of two copper(I) ions at the Cu_B site (Chang *et al.*, 2021; Chan *et al.*, 2021).

1.5.4 pMMO kinetics

The kinetic parameters of pMMO vary across the literature. They appear to be highly dependent on the growth conditions used and purification protocol, as well as the reductant used in the assay (generally formate for whole cells, NADH or duroquinol for membrane fractions, and duroquinol for solubilized or purified pMMO), and the organism used. The specific activity of pMMO has been calculated for methane oxidation, for example by measuring the rate of consumption of methane using gas chromatography (Lontoh and Semrau, 1998) or formation of methanol (Miyaji *et al.*, 2019). However, most studies carry out a propylene epoxidation assay to measure activity of pMMO, due to the product (propylene epoxide) not being further metabolised by the cells and being easily identified.

Methanotrophs expressing pMMO develop extensive intracytoplasmic membranes (ICMs), where pMMO is localised (Figure 1.6). ICMs contain large amounts of pMMO (up to 80% of total protein) (Lawton and Rosenzweig, 2016). Often ICMs are obtained from the cells by ultracentrifugation are used to perform experiments *in vitro* of membrane-bound pMMO. The enzyme has also been further solubilized or purified for example to obtain crystal structures. However, enzymatic activity of membrane-bound pMMO and especially of solubilized and purified enzyme is much lower than activity obtained for whole cells (Table 1.2).

Methanotrophs inhabit environments with high methane concentrations, such as paddy fields and landfill sites, where methane concentration is hundreds of times higher than the atmospheric value, and have apparent half-saturation constants (K_m) between 1–62 μM (Sirajuddin and Rosenzweig, 2015). However, one species has been found to express also a high affinity pMMO with a considerably lower K_m for methane of 0.11 μM which is capable of metabolizing atmospheric methane (Baani and Liesack, 2008). pMMO has a much narrower

substrate specificity than sMMO, with the ability to oxidise C1-C5 alkanes, and few alkenes and halogenated hydrocarbons (Sirajuddin and Rosenzweig, 2015).



Figure 1.6. ICM formation in methanotrophs expressing pMMO.

(A) Transmission electron micrograph of a *Methylosinus trichosporium* OB3b cell showing intracytoplasmic membranes (ICMs). Image from ('Diamond Light Source', 2015). (B) Transmission electron micrograph of *Methylococcus capsulatus* Bath grown with copper showing ICMs (Choi *et al.*, 2003).

Table 1.2. Kinetic parameters of pMMO.

Examples of specific activity reported in the literature for *Methylococcus capsulatus* Bath and *Methylosinus trichosporium* OB3b, for the oxidation of methane or propylene, in samples containing whole cells, isolated membranes containing, or solubilized pMMO. The turnover number per active site was calculated assuming that 20% of total protein is pMMO in whole cell samples, and that 80% of total protein is pMMO in membrane samples (as calculated in (Sirajuddin and Rosenzweig, 2015)), and using the molecular weight of 1 active site of 100 kDa. References: (1) (Sirajuddin *et al.*, 2014), (2) (Choi *et al.*, 2003), (3) (Lieberman *et al.*, 2003), (4) (Lontoh and Semrau, 1998), (5) (Miyaji *et al.*, 2019), (6) (Hakemian *et al.*, 2008).

Organism (substrate)	Sample type	Specific activity (nmol min ⁻¹ mg ⁻¹)	Turnover number per active site (s ⁻¹)
<i>Methylococcus capsulatus</i> Bath (methane)	Membrane	25–130 ⁽¹⁾	0.05–0.27
<i>Methylococcus capsulatus</i> Bath (propylene)	Whole cells	235 ⁽²⁾	1.96
	Membrane	16–81 ^(2,3)	0.03–0.17
	Solubilized or purified	3.9–17.7 ⁽³⁾	0.007–0.03
<i>Methylosinus trichosporium</i> OB3b (methane)	Whole cells	82–300 ⁽⁴⁾	0.68–2.5
	Membrane	4–6 ⁽⁵⁾	0.008–0.013
<i>Methylosinus trichosporium</i> OB3b (propylene)	Membrane	3 ⁽⁶⁾	0.006
	Solubilized or purified	0.11–2.9 ⁽⁶⁾	1.8x10 ⁻⁴ –0.005

1.6 MMOs in biotechnology

1.6.1 Methane bioremediation

MMOs are of interest for different applications in biotechnology. Firstly, they have been studied as catalysts to convert methane into biofuels and commercial chemicals due to their ability to break the C-H bond in methane at ambient temperature and pressure (Lawton and Rosenzweig, 2016; E. Bjorck *et al.*, 2018). Other studies have focused on the use of methanotrophs for the bioremediation of pollutants from soils such as trichloroethylene (TCE), lindane, and heavy metals, amongst others (Lee *et al.*, 2006; Semrau, 2011; Singh and Singh, 2017). Finally, methanotrophs have been studied for applications in biotic methane mitigation. For example, in landfills, biotarps and biocovers have been employed to significantly reduce emissions, consisting of the addition of coarse and porous materials which are rich in organic matter in a design that favours the growth of methanotrophs (Huber-Humer *et al.*, 2008). These systems have been employed to cover landfill sites in Austria obtaining methane removal rates of almost 100% in some cases (Huber-Humer *et al.*, 2008). However, these systems are not applicable to mitigation in natural wetlands and paddy fields. Moreover, the slow growth of methanotrophs and the need for high methane concentrations for their growth has led to attempts to obtain heterologous expression of MMOs in more amenable hosts such as *E. coli*. Nevertheless, in spite of attempts at heterologous expression, neither sMMO nor pMMO has been expressed in *E. coli*. The hydroxylase component of sMMO (MMOH) has been expressed in *Pseudomonas*, although methane oxidation activity of the recombinant strain was not shown (Koo and Rosenzweig, 2021). A technical report published by the Defense Technical Information Centre (DTIC) in the US describes the heterologous expression of sMMO in *Arabidopsis thaliana* plants, as a means to remove atmospheric methane (Strand *et al.*, 2013). Due to the inability of methanotrophs to grow at atmospheric methane concentrations, the authors suggest that pMMO expressed in plants to a high biomass could produce metabolization of atmospheric methane. Strand *et al.* (2013) achieved the transformation of

Arabidopsis thaliana with 5 genes: the three hydroxylase genes (*mmoX*, *mmoY*, *mmoZ*), *mmoC*, and *mmoG*, after trying different gene stacking methods. They report that transformant *Arabidopsis* plants have a functional sMMO, as evidenced by standard sMMO activity assays which detect the oxidation of naphthalene to naphthol (Strand *et al.*, 2013). The report also states that these plants have a “slight” methane removal capacity although further controls are needed. It is discussed that this potential low methane removal activity could be due to the absence of the regulatory protein B, with which they encountered problems during gene synthesis, and which has been shown to significantly enhance sMMO activity in methanotrophs (Mitić *et al.*, 2008). This report has thus set the grounds for the development of further research on the phytoremediation of methane using methanotroph catalysts, and highlighted some of the issues encountered, such as difficulties with carrying out effective gene stacking in plants and challenges with obtaining significant activity of sMMO in a heterologous host. Possibly parallel initiatives with pMMO comprised of only three genes could circumvent some of these issues.

1.6.2 Recombinant expression of pMMO

The recombinant expression of pMMO in a non-methanotroph organism is well recognised in previous literature as a challenge, principally due to the requirement of correct membrane insertion and sufficient copper content (Lawton and Rosenzweig, 2016). Recombinant expression in *E. coli* has not been achieved to date as pMMO genes appear to cause toxicity (Murrell *et al.*, 2000). Only recently, one study has managed to express the full-length pmoB subunit and truncated variants in *E. coli* K12 TB1 cells, a copper tolerant strain (Lu *et al.*, 2019). Neither the full-length subunit or the truncated versions exhibited methane or propylene oxidation activities, supporting the suggestion that the active site of the enzyme is not contained solely in pmoB (Lu *et al.*, 2019). One study has reported the heterologous expression of pMMO in *Rhodococcus erythropolis*, as a potential application to recycle biomass to methanol in a more industrially amenable host, as opposed to methanotrophs which present very slow growth

rates (Gou *et al.*, 2006). They cloned the enzyme under the promoter of a native *Rhodococcus* monooxygenase using a strain that is capable of growth on ethanol as a sole source of carbon, but not methane or ethane. As pMMO is able to convert ethane to ethanol this worked as a selection process for transformants. Gou *et al.* (2006) analysed the transcription and activity of cells expressing pMMO using both methane and propylene assays, and they reported a methane oxidation activity of $5.7 \mu\text{mol gDW}^{-1} \text{h}^{-1}$ of transformant *Rhodococcus* cells, compared to an activity of $1922 \mu\text{mol gDW}^{-1} \text{h}^{-1}$ of control *Methylosinus trichosporium* OB3b cells (Gou *et al.*, 2006). The activity of transformants was thus substantially lower than that of the methanotroph, however, the authors discuss that this could arise from the type of promoter used, or a lack of structural analysis to confirm the correct quaternary structure of the enzyme. The study also did not confirm membrane localisation of the complex, which is considered essential for its correct functioning, given the dramatic decrease or loss of activity of pMMO upon solubilisation. Nevertheless, the finding that cells were functionally expressing pMMO, albeit at low levels, demonstrates that the enzyme can function in a prokaryote with a different physiology and cell morphology to its original host (for instance *Rhodococcus* cells do not produce ICMs).

The recombinant expression of pMMO in plant cells has not been previously reported. Plant cells contain several membrane compartments which could potentially be suitable surfaces for pMMO expression and assembly. In the case of achieving expression of active pMMO in plant cells, the downstream metabolic steps following methane oxidation could be investigated regarding the production and removal of methanol. Methanol is naturally produced in plants during the degradation of pectin methyl esters in the cell wall and released through the stomata (McGiffen and Manthey, 1996; Yurimoto *et al.*, 2021). Although plants do not possess known enzymes which catalyze the oxidation of methanol, some studies have found using labelled methanol that the alcohol is partly converted to CO_2 and organic molecules in plant species (Cossins, 1964; Gout *et al.*, 2000). Moreover, other studies have found that plants are resilient to certain levels of methanol (Ramírez *et al.*, 2006). Therefore, it may be the case that plants

expressing active pMMO would not suffer from methanol toxicity. Alternatively, strategies for methanol removal, such as favouring the growth of methanol-utilizing bacteria and yeasts present on the leaf surface could be pursued.

1.7 Objectives and approach

The aim of this work was to recombinantly express pMMO in plants for methane detoxification. The work was carried out primarily in *Nicotiana tabacum* (tobacco) and *Arabidopsis thaliana*, two model organisms, as a proof of concept, with the aim of ultimately creating transgenic rice for methane mitigation of paddy fields. pMMO was chosen as a catalyst over sMMO as pMMO is ubiquitous in nature and functional in many environments and varying copper levels, as well as having a lower substrate specificity, which could lower the amounts of unwanted chemical reactions taking place in the recombinant host. Moreover, pMMO is encoded by only three genes, as opposed to six genes encoding sMMO plus an additional two genes required for correct transcription and assembly, rendering pMMO a more methodologically feasible target with which to create transgenic plants. Due to pMMO being an integral membrane protein expressed in the stacked ICMs in methanotrophs, the plant endoplasmic reticulum (ER) was considered a suitable membrane compartment in which to house the recombinant enzyme complex.

The objectives were:

- To transiently express the pmoC, pmoA and pmoB subunits in tobacco ER, using fluorescent tags to determine localization, and to determine pairwise protein-protein interactions, and confirm topology of the membrane proteins (Chapter 3)
- To obtain stoichiometric co-expression of all three subunits in tobacco using an appropriate gene stacking method which allows the co-ordinated expression of the subunits and ER targeting (Chapter 4)

- To quantify the impact of the expression of a pMMO polypeptide on ER structure and dynamics (Chapter 5)
- To obtain stable expression of the pMMO polypeptide in rice plants and Arabidopsis and determine enzymatic activity of the complex (Chapter 6).

2 Materials and methods

2.1 Plant material and growth conditions

Tobacco

Experiments were performed using wild type *Nicotiana tabacum* (SR1 cv Petit Havana), referred to as tobacco in this study, or *Nicotiana benthamiana* if specified. Seeds were placed on sterile ½ MS plates (2.4 gL⁻¹ Murashige and Skoog medium, 0.8% phytoagar in deionised water, pH 5.7), sealed with a double layer of micropore tape, and placed in tobacco growth chambers (23/20°C, 12 hs light/dark). After 10 days, seedlings were transferred to soil and grown in greenhouse conditions (21°C, 14 hours light, 10 hours dark). 5-6 week old plants were used for agroinfiltration.

Arabidopsis

Wild type *Arabidopsis thaliana* (Colombia-0) seeds, referred to hereafter as Arabidopsis, were sterilised using 70% ethanol and mixing by inversion for 3 minutes. Seeds were placed in sterile ½ MS plates and sealed with a double layer of micropore tape. Plates were then covered in foil and placed at 4°C for three days for stratification. Plates were then transferred to Arabidopsis growth cabinets (22°C, 16 hours light, 8 hours dark) for 10 days, after which seedlings were transferred to soil and grown in greenhouse conditions (21°C, 14 hours light, 10 hours dark).

Rice

Rice plants (*Oryza sativa* ssp. *japonica* cv. Nipponbare) were transformed at the National Institute of Agricultural Botany (NIAB) and provided to Oxford Brookes University as young seedlings in Jiffy Pellets. Plants were transferred to pots containing 100% clay granules (Profile “Greens”), and kept in flooded conditions by adding warm water (approximately 25°C) to trays. Fertiliser solution (Peter’s Excel fertiliser N.P.K-15-5-15) was added once a week (240gL⁻¹

stock solution diluted 1:100), and a sequestered iron solution (Chempak) was added to the fertiliser every other week (48gL⁻¹ stock solution diluted 1:100) (adapted from (McCouch RiceLab, Cornell University)). Iron solution was not added after flowering. Plants were grown in long day tropical greenhouse conditions (30°C, 16 hours light, approximately 60% humidity) until the maximum tillering stage and transferred to incubators in short day conditions for flowering (28/25°C day/night temperatures, 12 hours light).

2.2 Cloning

2.2.1 Gateway cloning

Expression plasmids were constructed using Gateway® Technology, which is based on the bacteriophage lambda site-specific recombination system, and allows the incorporation of a gene of interest into vector by recombination in *att* sites. Genes were amplified by PCR using Q5 high-fidelity DNA polymerase from New England Biolabs (NEB) adding *attB* sites as flanking sequences. Primers were synthesized by Eurofins Genomics. 50 µl PCR reactions were run on a 0.7% agarose gel, bands were excised with a clean scalpel and the DNA was extracted using the NEB Gel Extraction kit. PCR fragments were added to a BP reaction following the manufacturer's instructions using half the volume, for incorporation of the gene of interest to a pDONR vector (pDONR221 or pDONR207), yielding an entry clone. 1 µl of BP reaction was used to transform NEB 5-alpha Competent *E. coli* (High efficiency) by heat shock, following manufacturer's instructions using half of the volume of recommended cells and a fifth of the volume of recommended SOC culture media. Cells were grown overnight on LB agar plates containing the appropriate antibiotic as selection agent.

Colonies were tested for the presence of transgenes by PCR using OneTaq DNA Polymerase (NEB), and positive clones were grown overnight in LB liquid culture at 37°C and shaking at 200 rpm. Plasmid DNA was purified using the NEB Monarch Plasmid Miniprep Kit and confirmed by sequencing (Eurofins Genomics). Subsequently, purified entry clones were

added to an LR reaction, in which recombination with a destination vector yields the final expression plasmid. Following the LR reaction, competent *E. coli* were transformed and expression plasmids were purified as described previously for the BP reaction. Expression plasmids contain a left border (LB) and a right border (RB) delimiting the T-DNA for integration into the plant nuclear genome, as well as bacterial and plant resistance genes to enable selection. *E. coli* selections were carried out using 50 $\mu\text{g mL}^{-1}$ kanamycin or spectinomycin. The expression plasmids were used for *Agrobacterium* transformation as described in section 2.3.

2.2.2 Cloning of single subunits with fluorescent tags

Expression plasmids containing the pMMO subunits fused to fluorescent proteins were created with the Gateway cloning method described above. The DNA sequences of the pMMO genes were previously amplified by PCR by Verena Kriechbaumer from genomic DNA of *Methylosinus trichosporium* OB3b (accession number: U31650.2), kindly provided by Professor Tom Smith, Sheffield Hallam University. pDONR vectors containing pmoC, pmoA and pmoB without stop codons were supplied by Verena Kriechbaumer and used either directly for LR reactions into destination vectors, or used as a source of DNA for cloning N-terminally tagged constructs and pmoB variants.

The expression plasmids containing single subunits fused to fluorescent tags under a strong constitutive promoter (35S) created in this study are listed in Table 2.1.

Table 2.1. List of single subunit expression plasmids.

The following plasmids were created in this study for use in subcellular localization analysis, targeting pmoB to the ER, FRET-FLIM and redox-based topology analysis. The vector backbones pSS01 and pCM01 are described in (Brach *et al.*, 2009), the remaining vectors derive from (Karimi *et al.*, 2002).

Expression plasmid	Vector backbone
GFP-pmoC	pB7WGF2
GFP-pmoA	pB7WGF2
GFP-pmoB	pB7WGF2
pmoC-GFP	pB7FWG2
pmoA-GFP	pB7FWG2
pmoB-GFP	pB7FWG2
Btrunc0-GFP	pB7FWG2
Btrunc1-GFP	pB7FWG2
Btrunc2-GFP	pB7FWG2
Btrunc3-GFP	pB7FWG2
spBtrunc-GFP	pB7FWG2
spBtrunc-GFPhdel	pB7FWG2
spBtrunc-GFPhdel2	pB7FWG2
spBtrunc-RFP	pB7RWG2
spBtrunc-RFPhdel	pB7RWG2
spBtrunc-RFP-KKRY	pB7RWG2
pmoC-RFP	pB7RWG2
pmoA-RFP	pB7RWG2
CFP-pmoA	pK7WGC2
roGFP2-pmoC	pCM01
roGFP2-pmoA	pCM01
pmoC-roGFP2	pSS01

pmoA-roGFP2	pSS01
spBtrunc-roGFP2	pSS01

The primers designed for the construction of the vectors listed above are shown in Table 2.2. pmoA_{gatefor}/rev and pmoC_{gatefor}/rev were used for cloning pmoA and pmoC in pDONR221 with stop codons for N-terminally fused fluorophores. PmoB signal peptide truncations (0-3) were constructed using primers fBtrunc (0-3) each with rpmoB392. The plant signal peptide used to create spBtrunc-GFP comes from pumpkin 2S albumin (amino acid sequence: MARLTSIIALFAVALLVADAYAYRT) (Hara-Hishimura *et al.*, 1993) and was added via a nested PCR, using spBtrunc1 forward primer initially, and subsequently the gatesp forward primer on the PCR product to add the Gateway overhang. The C-terminal additions to pmoB (HDEL, KKRY and HDEL2 sequences) were included in the reverse primers, and PCR was carried out using spBtrunc-GFP or spBtrunc-RFP expression plasmids as the DNA source. Primers f/rpmoC207, f/rpmoA207 and f/rpmoB207 were used for cloning the subunits into pDONR207 for LR reaction into the pSS01 vector for redox-based topology analysis.

Table 2.2. List of primers for construction of single subunit expression plasmids.

Primer sequences designed for the construction of vectors in Table 2.1. Gateway overhangs are highlighted in grey.

Primer name		Sequence (5'-3')
pmoAgatefor	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTTA ATGTTTACATCGAAGAGC
pmoAgaterev	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGT TCAGATCTTCTGGATCAC
pmoCgatefor	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTTA ATGAGCGTAACAACAGA
pmoCgaterev	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGT TCAGCCGACCAG
fBtrunc0	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTCT ATGGCGTTTCTGCGCATG
fBtrunc1	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTCT ATGCACGGCGAGAAGTCGC
fBruntc2	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTCT ATGTCGGTGGCCCCGG
fBtrunc3	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTCT ATGGGCCTGAGCGTTGC
spBtrunc1	Forward	ATGGCCAGACTCACAAGCATCATTGCCCTCTTCGCAGTG GCTCTGCTGGTTGCAGATGCGTACGCCTACCGCACT CACGGCGAGAAGTCGC
gatesp	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTCT ATGGCCAGACTCACAAGCA
rpmoB392	Reverse	GGGGACCACTTTGTACAAGAAAGCT GGGTAGGGCATGTCGCCG
GFPhdelr	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTA TTACAGTTCATCATGCTTGTACAGCTCGTCCATGCCG
GFPhdelr2	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTA TTAAAGCTCATCGTGGTGGTGGTGGTGCCCCCCCC CTTGTACAGCTCGTCCATGCCG
RFPhdelr	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTA TTACAGTTCATCATGGGCGCCGGTGGAGTGG
RFPkkr	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTA TTAATATCTCTTCTTGCGCCGGTGGAGTGG
fpmoC207	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCT ATGAGCGTAACAACAGAGAC
rpmoC207	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGT GCCGACCAGGGCC
fpmoA207	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCT ATGTTTACATCGAAGAGCGG
rpmoA207	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGT GATCTTCTGGATCACCTTGG
fpmoB207	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCT ATGGCCAGACTCACAAG
rpmoB207	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGT GGGCATGTCGCCG

Additional expression plasmids used in this study as marker proteins or for other analysis with the single subunits which were constructed previously are listed below, specifying the vector of origin, the contributor to their creation and the source publication when present (Table 2.3).

Table 2.3. Expression plasmids constructed previously.

Expression plasmids containing organelle marker proteins and pMMO subunits used in this study, specifying the contributor to their creation and source publication.

Expression plasmid	Vector backbone	Made by	Publication
RFP-HDEL	pVKH18En6	Lee <i>et al</i>	(Lee <i>et al.</i> , 2013)
GFP-HDEL	pVKH18En6	Federica Brandizzi	(Brandizzi <i>et al.</i> , 2003)
GFP-CNX	pVKH18En6	Sarah Irons	(Irons <i>et al.</i> , 2003)
ST-RFP	pVKH18En6	Federica Brandizzi	(Renna <i>et al.</i> , 2005)
RFP-ATPase	pDCL	Candat <i>et al</i>	(Candat <i>et al.</i> , 2014)
mCherry-pmoC	Modified pB7WGF2	Verena Kriechbaumer	Unpublished
Clover-pmoA	Modified pB7WGF2	Verena Kriechbaumer	Unpublished
roGFP2	pEarleyGate100	Brach <i>et al</i>	(Brach <i>et al.</i> , 2009)
roGFP2-HDEL	pWEN22	Brach <i>et al</i>	(Brach <i>et al.</i> , 2009)

2.2.3 Construction of 2in1 vectors

2in1 destination vectors were kindly provided by Christopher Grefen (Hecker *et al.*, 2015). The vectors contain two cassettes flanked by *att* regions within a same T-DNA, producing two separate fusion proteins, as illustrated in Figure 2.1. Each cassette is placed under a 35S promoter and omega leader sequence. Different destination vectors allow in frame fusion to mCherry and EGFP fluorophores, in the N-terminus, C-terminus or combination of both. Six 2in1 vectors were constructed varying the subunit in each cassette, the combination of subunits

and the position of the fluorophore (Table 2.4). The primers used for cloning are listed in Table 2.5.

The cloning procedure was as described above, employing pDONR221-P1P4 for integration into the first cassette, and pDONR221-P3P2 for integration into the second cassette. The LR reaction contained both pENTRY vectors, alongside a single destination vector. Following *E. coli* transformation with the LR reaction, integration into the first cassette was selected by absence of the *ccdB* gene (as with single cassette Gateway vectors), and integration into the second cassette was selected through blue/white screening, with white colonies indicating the absence of the *lacZ* gene. Agar plates contained spectinomycin, 40 $\mu\text{g mL}^{-1}$ X-gal and 0.1mM IPTG (for blue/white screening). Stock solution of X-gal was made in DMSO and covered in foil for storage at -20°C . IPTG solution was made in water and stored at -20°C .

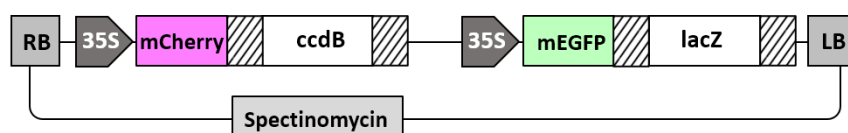


Figure 2.1. Schematic representation of 2in1 vectors.

Example of a 2in1 destination vector for N-terminal fusion of mCherry and EGFP. RB= right border, LB= left border, 35S= Cauliflower mosaic virus 35S promoter, *ccdB*= gyrase inhibitor gene, *lacZ*= b-galactosidase promoter and α -subunit of b-galactosidase, spectinomycin= aminoglycoside-39-adenylyltransferase expression cassette. Striped rectangles indicate *att* Gateway recombination sites. Modified from (Hecker *et al.*, 2015).

Table 2.4. List of 2in1 expression vectors.

Enumeration of 2in1 expression vectors constructed in a combinatorial manner, varying the subunit in each cassette, the combination of subunits and the position of the fluorophore (named as first cassette/ second cassette).

2in1	Vector backbone
1. mCherry-pmoC/ EGFP-pmoA	pFRETgc
2. mCherry-pmoB/ EGFP-pmoA	pFRETgc
3. pmoC-mCherry/ pmoA-EGFP	pFRETgc
4. pmoB-mCherry/ pmoA-EGFP	pFRETgc
5. mCherry-pmoC/ pmoA-EGFP	pFRETgc
6. pmoA-mCherry/ pmoC-EGFP	pFRETgc

Table 2.5. Primers for construction of 2in1 vectors

Primer names indicate forward or reverse (f/r), in the case of reverse primers, the next letter C/N states if the fluorophore is at the C-terminus (no stop codon) or at the N-terminus (with stop codon), the next letter in both forward and reverse primers shows the pMMO subunit (C, A or B), followed by the distinction between 1st and 2nd cassette (Gateway overhangs slightly differ). As an example rNC1st means reverse primer, N-terminal tag, pmoC subunit, 1st cassette. Primers were designed according to a protocol from the Grefen lab, University of Tübingen. Gateway overhangs are highlighted in grey.

Primer name		Sequence (5'-3')
fC1st	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGC TTAATGAGCGTAACAACAGAGAC
rNC1st	Reverse	GGGGACAACCTTTGTATAGAAAAGTTGGGT TCAGCCGACCAGGG
rCC1st	Reverse	GGGGACAACCTTTGTATAGAAAAGTTGGGTG GCCGACCAGGGC
fA2nd	Forward	GGGGACAACCTTTGTATAATAAAGTTGGA ATGTTTACATCGAAGAGCGG
rNA2nd	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGT TCAGATCTTCTGGATCACCTTG
rCA2nd	Reverse	GGGGACCACTTTGTACAAGAAAGCTGG GTGGATCTTCTGGATCACCTTGG
fA1st	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGC TTAATGTTTACATCGAAGAGCGG
rNA1st	Reverse	GGGGACAACCTTTGTATAGAAAAGTTGGGT TCAGATCTTCTGGATCACCTTG
fB2nd	Forward	GGGGACAACCTTTGTATAATAAAGTTGGA ATGAAAGCTCTGGAAAGAATGG

rNB2nd	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGT TCAGGGCATGTCGC
fB1st	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGC TTAATGAAAGCTCTGGAAAGAATGG
rCB1st	Reverse	GGGGACCACTTTGTATAGAAAAGTTGGGT GGGGCATGTCGCCG
rCA1st	Reverse	GGGGACAACCTTTGTATAGAAAAGTTGGGT GGATCTTCTGGATCACCTTG
fC2nd	Forward	GGGGACAACCTTTGTATAATAAAGTTGGA ATGAGCGTAAC AACAGAGAC
rCC2nd	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGT GGCCGACCAGGGC

2.2.4 Cloning of polypeptide vectors

The polypeptide sequence containing all pMMO subunits and clover separated by 2A self-cleaving peptides was designed following the reports in (Liu *et al.*, 2017) regarding the choice of self-cleaving peptides and their order in quad-cistronic vectors, and using the ER-targeted pmoB (spBtrunc). The sequence was codon optimized and synthesized by Twist Bioscience. The DNA and protein sequence are shown in Appendix 1. Primers were designed with Gateway overhangs to clone the full polypeptide and polypeptides containing less genes into pDONR221 (2A self-cleaving peptides are shown in bold):

“Polypeptide”: pmoC-**T2A**-pmoA-**P2A**-spBtrunc-**E2A**-clover

“CAB”: pmoC-**T2A**-pmoA-**P2A**-spBtrunc

“CA”: pmoC-**T2A**-pmoA

LR reactions were performed to obtain N-terminally RFP-tagged constructs (Table 2.6). Two of the expression vectors generated (RFP-CAB and RFP-CA) were used as a DNA source to add a C-terminal GFP, by BP reaction into pDONR221 and LR reaction into a C-terminal GFP destination vector, generating “double fluorophore fusion constructs”: RFP-CAB-GFP and RFP-CA-GFP.

A second polypeptide referred to in this study as “polyMdomain” or “polyM” was designed with some modifications from the original polypeptide. Firstly, a self-excising intein domain was added preceding the P2A signal peptide (Zhang *et al.*, 2017), for excision of the amino

acids from P2A which would otherwise remain added to the C-terminus of pmoA. Secondly, the plant signal peptide used to direct pmoB to the ER was from tobacco pathogenesis related protein PR1a (amino acid sequence: MGFVLFSQLPSFLLVSTLLLFLVISHSCRA), to avoid the addition of three N-terminal amino acids to pmoB as was the case with the pumpkin 2S albumin signal peptide. Thirdly, a domain with multiple N-glycosylation sites (“M domain”) was added following the pmoB subunit, which reportedly increases expression levels of ER proteins in plants (Kang *et al.*, 2018). The full DNA and protein sequence of polyM are shown in Appendix 2.

“PolyM”: pmoC-T2A-pmoA-InteinP2A-PR1Btrunc-Mdomain

The polyM sequence was codon optimized and synthesized by Twist bioscience, and the pTwist ENTRY clone was used in an LR reaction for N-terminal RFP tagging and C-terminal GFP tagging.

Table 2.6. List of polypeptide expression vectors.

Polypeptide expression vectors constructed in this study. Vector backbones from (Karimi *et al.*, 2002).

Expression vector	Vector backbone
1. RFP-polypeptide	pK7WGR2
2. RFP-CAB	pK7WGR2
3. RFP-CA	pK7WGR2
4. RFP-CAB-GFP	pB7FWG2
5. RFP-CA-GFP	pB7FWG2
6. RFP-polyM	pK7WGR2
7. polyM-GFP	pB7FWG2

The primers used for the construction of polypeptide vectors are shown in Table 2.7: GatepolyCf with Gatepolycloverr/GateplyBr/GatepolyAr were used for cloning the polypeptide, CAB and CA into pDONR221 and creating RFP-tagged constructs, the remaining

primers were used for cloning from the RFP-tagged expression clones and creating the double fluorophore fusions.

Table 2.7. Primers used for polypeptide vector construction.

List of primers used for the cloning of polypeptide vectors. Gateway overhangs are highlighted in grey.

Primer name		Sequence (5'-3')
GatepolyCf	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTCC ATGTCTGTCACAACTGAAA
Gatepolycloverr	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTA CTACTTATAGAGCTCATCCA
GatepolyBr	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTA CTAAGGCATGTCTCCAGC
GatepolyAr	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTA CTATATTTTTTGTATCACTTTGGT
GateRFPf	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTCC ATGGCCTCCTCCGAG
polyArNS	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTA TATTTTTTGTATCACTTTGGTGGT
polyBrNS	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTA AGGCATGTCTCCAGC

2.3 Transformation of competent *Agrobacterium* cells

Competent *Agrobacterium tumefaciens* cells (strain GV3101), referred to hereafter as *Agrobacterium* cells, were made starting with a 5ml pre-culture of cells in LB media incubated overnight at 28°C and 180 rpm, with 25 µg mL⁻¹ rifampicin. This was employed to inoculate a larger volume (200 ml) of LB in a sterile beaker, which was incubated overnight in the same conditions, and grown to an OD₆₀₀ of 0.5 to 1. The culture was split into 50 ml falcon tubes, chilled on ice for 5 minutes, and centrifuged at 3500 rpm for 30 minutes at 4°C. Pellets were resuspended in 15 ml of ice cold CaCl₂ solution (20mM, 15% glycerol), then spun at 5000 rpm for 10 minutes at 4°C, and again resuspended in 15 ml of ice cold CaCl₂ solution. The competent cells were aliquoted, snap frozen in liquid N₂ and stored at -80°C.

For transformation, cells were thawed on ice and separated into 50 μ l aliquots, and approximately 250 ng of DNA (expression vectors) was added to each aliquot. Cells were incubated on ice for 5 minutes, subsequently at -80°C for 3 minutes, and then incubated in a 37°C water bath for 4 minutes, to allow the adsorbed DNA to enter the cells by heat shock. Following heat shock, 1 ml of LB media was added and cells were allowed to recover at 28°C whilst shaking for 2 hours. Cells were spread on agar plates containing $25\text{ }\mu\text{g ml}^{-1}$ rifampicin and the appropriate plasmid selection agent. Plates were incubated at 28°C for two days for the formation of colonies. Single colonies were grown in 3 ml liquid cultures overnight at 28°C and 180 rpm, with $25\text{ }\mu\text{g ml}^{-1}$ rifampicin plus the plasmid selection agent, for their use in the transient transformation of tobacco, or for their storage in glycerol stocks at -80°C (200 μ L of 80% sterile glycerol plus 800 μ l of bacterial suspension).

2.4 Transient transformation of tobacco

Tobacco plants were transformed by agroinfiltration following a procedure adapted from (Sparkes *et al.*, 2006). 2 mL of overnight agrobacterium liquid cultures were spun down at 4000 rpm for 5 minutes and gently resuspended by pipetting in 1 ml of infiltration buffer (5 mg mL^{-1} glucose, 50 mM MES, 2 mM $\text{Na}_3\text{PO}_4 \cdot 12\text{H}_2\text{O}$ and 0.1 mM acetosyringone). A second wash step was performed by repeating the centrifugation and resuspension in 1 mL of fresh infiltration buffer, these wash steps remove the antibiotics from the media. Cultures were diluted to an appropriate OD_{600} for infiltration (0.1 unless otherwise stated) with infiltration buffer.

Greenhouse grown plants, approximately 5-6 weeks old, were used for transformation. A hypodermic needle was used to make a small puncture on the abaxial surface of the leaves. Using a 1 mL syringe the cultures were infiltrated into the leaves by slowly discharging the suspension on the puncture whilst holding the upper side of the leaf. Once finalised, plants were watered and placed in tobacco growth chamber until imaged by confocal microscopy

(generally 48 hours after infiltration). For experiments with p19, *Nicotiana benthamiana* was used instead of tobacco, and p19 was co-infiltrated with the construct of interest at an OD₆₀₀ of 0.1.

2.5 Generation of stable Arabidopsis plants

Stable Arabidopsis plants were produced by the floral dipping method (Clough and Bent, 1998). Agrobacterium cultures were grown in LB containing antibiotics in 250 ml sterile flasks at 28 °C and 180 rpm overnight. Each culture was split into two 50 ml falcon tubes and centrifuged at 4000 rpm for 10 minutes. The pellets were resuspended in a 5% sucrose, 500 µl L⁻¹ Silwet 77 solution. Arabidopsis plants at flowering stage were immersed in the solution for 30 seconds with agitation and placed in a tray covered with cling film to maintain humidity for 24 hours. The plants were then watered and kept in greenhouse conditions and the floral dip was repeated a week later to increase the transformation efficiency. After a week of watering, the plants were bagged and left to dry for seed production. For selection, seeds were sterilised with 70% ethanol and mixing by inversion for 3 minutes, then placed on plates with MS media and 50 mg L⁻¹ kanamycin. Putative transformants (T0) were distinguished by their growth on the selective medium as opposed to untransformed seeds which were unable to grow. T0 plants were grown on plates until 10 days old and then transferred to soil and grown in greenhouse conditions. The presence of the transgene was checked by confocal microscopy. Plants were grown to maturity and bagged and dried for seed production (T1 generation). T1 seeds (each one either homozygous for the transgene, heterozygous or lacking the transgene) were grown on selective media as described above and each plant bagged individually for the isolation of T2 seeds to isolate homozygous lines.

2.6 Confocal microscopy

A small leaf piece (approximately 6 mm²) was cut with a scalpel and placed on a glass slide with the abaxial side facing up. Water was added and the slides were imaged on the abaxial side by confocal laser scanning microscopy (CLSM). In CLSM a laser beam is focused on a defined spot in the sample where it excites fluorescent molecules, and scans the sample in a raster pattern to achieve a 2D image. Confocal imaging is based on the use of spatial filtering to prevent out of focus light arising from different focal planes from reaching the detector, this is achieved by means of a pinhole, and greatly reduces the signal to noise ratio and enhances image resolution compared to widefield fluorescence microscopy.

Imaging was carried out at Oxford Brookes University using either a Zeiss LSM800 with an Airyscan detector or a Zeiss LSM880 with an Airyscan detector. The Zeiss LSM880 Airyscan detector consists of a 32-channel gallium arsenide phosphide (GaAsP) photomultiplier tube (PMT) detector, with 32 hexagonally arranged arrays each acting as individual very small pinholes and greatly increasing resolution (Huff, 2015). Images were taken on the LSM880 with a PlanApo 63x/1.46 NA oil immersion objective or a PlanApo 100x/1.46 NA oil immersion objective, or on the LSM800 on a PlanNeo 40x/1.30 NA oil immersion objective. The fluorophores used in this work and their excitation wavelengths were as follows: green fluorescent protein (GFP) excited at 488 nm, red fluorescent protein (RFP) excited at 561 nm, and cyan fluorescent protein (CFP) excited at 405 nm. The detection wavelengths were determined using the “smart set up” function in the Zeiss software. For multi-channel imaging a multi-track set up was used to reduce bleed-through between channels, and line switching for more accurate temporal correspondence between the channels. Laser power and gain were adjusted according to experiment needs and expression levels. Leaves were imaged at two days after infiltration, unless otherwise specified.

2.7 Quantifying fluorescence intensity levels

2.7.1 Experimental design

Following transient transformation of tobacco, individual cells present very variable expression levels, especially in low expressing constructs, where most cells have trace signal levels but a few cells in a leaf piece have relatively high fluorescence levels. Therefore, to quantify the expression levels of a construct, scanning individual cells would introduce errors due to the stochasticity of sampling. In order to reduce stochasticity, a quantification method was established based on scanning multiple ROIs with multiple cells each, in order to cover a larger amount of cells from the sample. Images were taken using the 40x objective using low magnification in order to include several cells within each ROI, and 20 ROIs (technical repeats) were imaged per leaf piece in order to scan a significant amount of area ($20 \times 200 \mu\text{m}^2$ ROIs = 67% of a 6 mm^2 leaf piece). The mean fluorescence intensity of the 20 ROIs was considered an estimate of the intensity level of that construct in a certain plant (biological replicate). All constructs were infiltrated at the same OD_{600} (0.1) on the same leaf within a biological replicate (as represented in Figure 2.2 with an example for 6 constructs). Generally three biological replicates were taken.

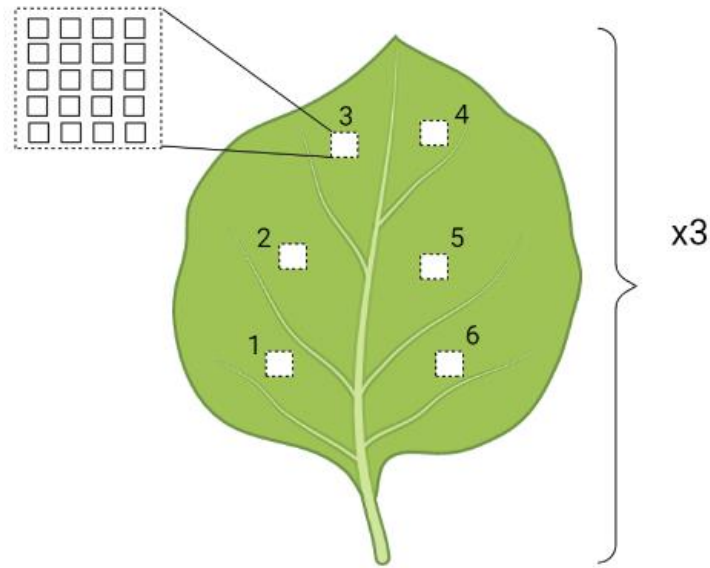


Figure 2.2. Experiment design for quantification of fluorescence intensity.

All constructs (1-6) were infiltrated at the same OD₆₀₀ on the same leaf within a biological replicate. A leaf piece from each construct was imaged (white squares), 20 ROIs were imaged from each leaf piece (inset left) to calculate a mean fluorescence intensity level. The experiment was repeated 3 times.

2.7.2 Image processing

Imaging was carried out on the Zeiss LSM 800. Due to the large difference in fluorescence intensity between the high and low expressing constructs, the laser power (lp) and gain were kept low (i.e. lp=0.6%, gain=600) to capture most of the dynamic range, that is, reducing the amount of saturated pixels in high expressing constructs and obtaining some detectable signal in the lower expressing constructs. The focal plane was chosen on each ROI as to maximize amount of ER surface shown. However, different amounts of ER area were captured in each ROI according to the flatness of the cells in that ROI, the size of cells, and number of cells expressing. Therefore, to eliminate the contribution of these variables to the fluorescence levels, an automated threshold was applied on each image (the "Default" threshold method in ImageJ described as a modified IsoData method), which binarizes the image into areas of background and signal, and subsequently the average pixel intensity was calculated on the thresholded area. Example images are shown in Figure 2.3 and the same images following processing and application of the threshold are shown in Figure 2.4.

When applying the threshold to low expressing constructs, such as spBtrunc and its variants, autofluorescence from stomata and the cell borders was often included in the thresholded area, due to the low contribution of the ER signal to the total signal. For this reason, the group “autofluorescence” was included in the comparison, imaged from leaf pieces which were not infiltrated. Given the low overall signal in the autofluorescence ROIs, on occasion, the threshold would select only very small regions of the guard cells yielding an abnormally high mean intensity measurement, therefore these images were removed (Figure 2.5).

Before applying a threshold, images were processed in ImageJ using background subtraction and a median filter (3x3 kernel) to reduce the influence of extreme pixel values. The processing and measurements were automated using modified ImageJ plugins to split channels, subtract background, apply a median filter, apply a threshold and measure the intensity of the thresholded area sequentially through a folder.

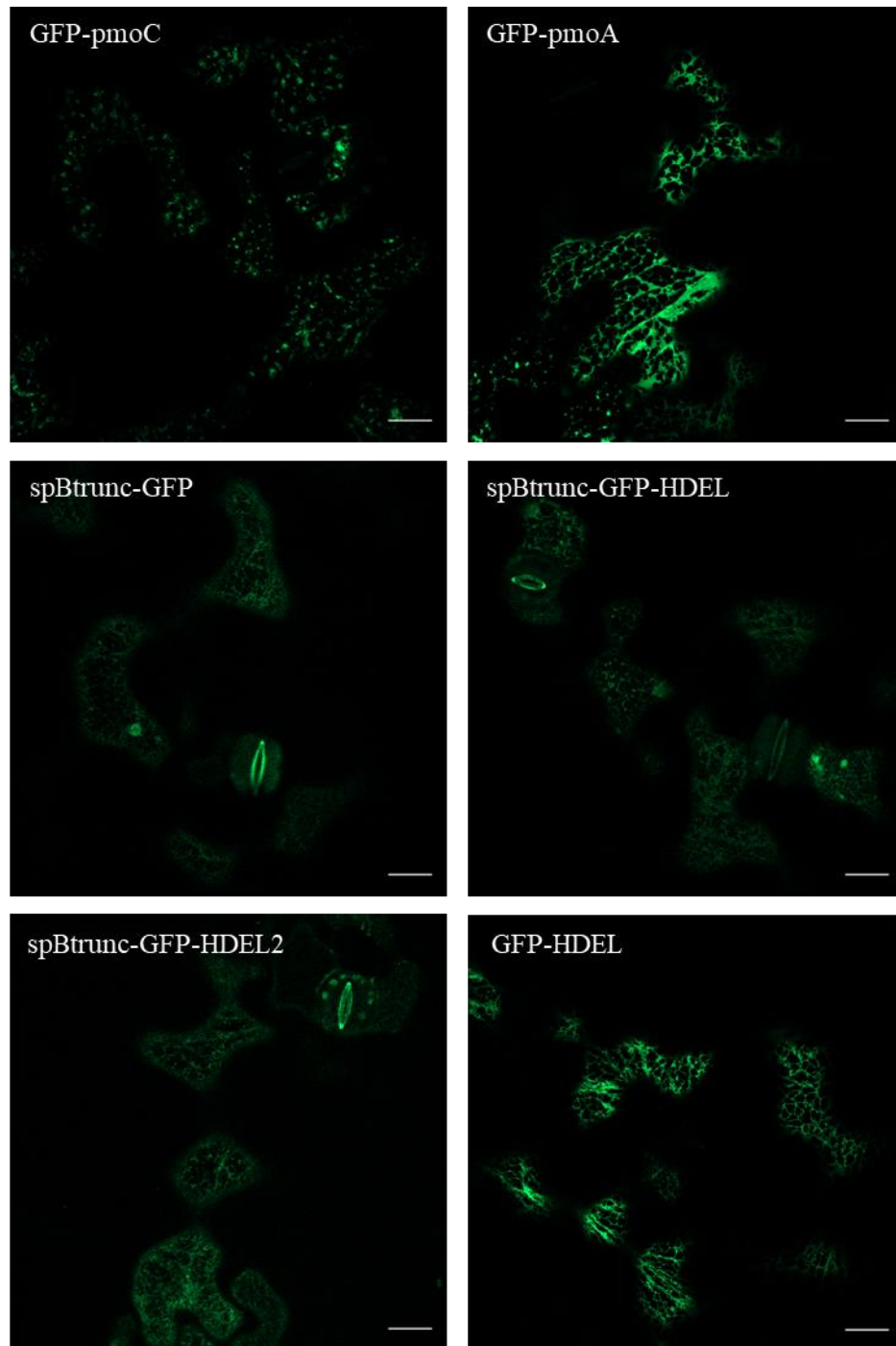


Figure 2.3. Example images for quantification of fluorescence intensity.

Example ROIs used for the quantification of signal intensity of GFP-tagged constructs at OD=0.1. Images taken using a 40x objective, 0.6% lp and gain 600. Scale bar = 20 μ m.

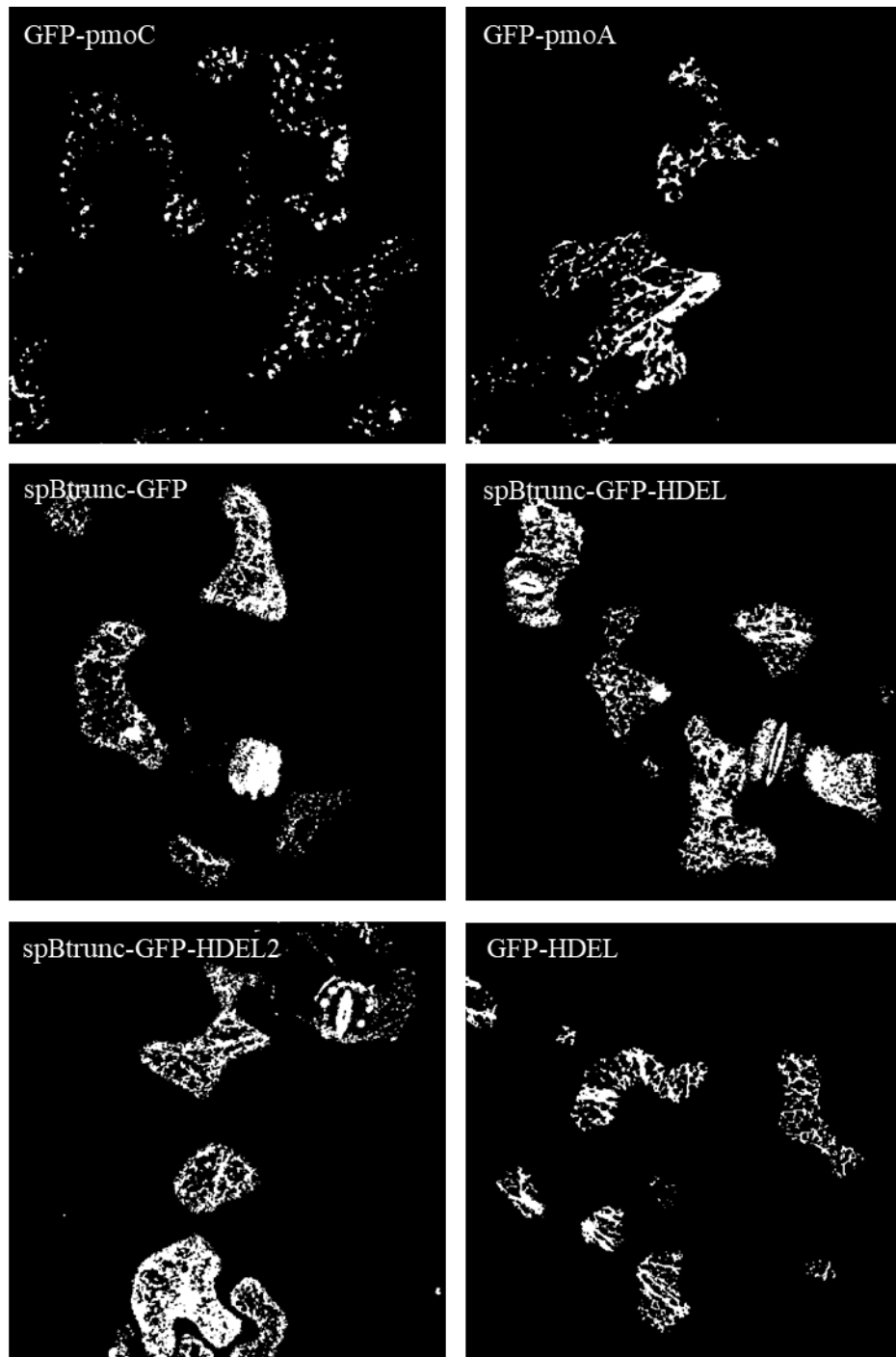


Figure 2.4. Application of a threshold.

Images following background subtraction, application of a median filter (3x3) and setting of an automated threshold. White areas represent areas above the threshold value which were considered for calculating the mean fluorescence intensity.

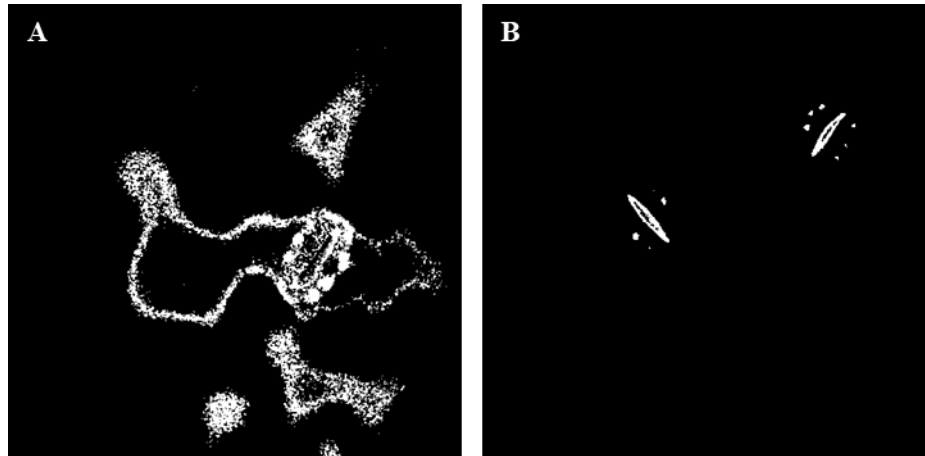


Figure 2.5. Manual curation of thresholded images.

(A) Threshold applied to an autofluorescence image which selected several cell borders and elements of the guard cells, and (B) threshold applied to an autofluorescence image which selected a very reduced region of the guard cells and chloroplasts, producing an abnormally high fluorescence reading – these images were removed from the dataset.

2.8 Redox ratio analysis

A redox-based ratio analysis was performed to determine ER membrane topology based on the use of roGFP2, a modified GFP protein which emits fluorescence with different efficiencies when excited at 405 nm and at 488 nm depending on its oxidation state (Hanson *et al.*, 2004; Brach *et al.*, 2009). Soluble cytosolic roGFP2 (“roCyto”) and the ER lumen localized “roGFP-HDEL” were used as references for both compartments. The following constructs were cloned (as described in Section 2.2.2): roGFP2-pmoC, roGFP2-pmoA, pmoC-roGFP2, pmoA-roGFP2, and spBtrunc-roGFP2. Constructs were infiltrated into *Nicotiana benthamiana* with p19, except for roCyto and roGFP-HDEL which were infiltrated into *Nicotiana benthamiana* without p19. Images were acquired with the Zeiss 880 PlanApo 63x/1.46 NA oil immersion objective, using a multitrack mode with line switching to measure the autofluorescence, excitation at 405 nm and excitation at 488 nm. Different imaging conditions were adjusted and tested with the roCyto and roGFP-HDEL constructs until the maximum possible 405/488 ratio separation between the controls was achieved. Given the substantially higher excitation efficiency of the 488 nm laser in all oxidation states, a large difference in the laser power was

required to obtain a significant signal in the 405 channel without saturation on the 488 channel. Additionally, conditions were adjusted to account for a large difference in expression levels between the controls and the pMMO constructs, the latter expressing at much lower levels. The optimised imaging conditions which were used for analysis are shown in Table 2.8.

Due to the much higher expression of control proteins compared to pMMO proteins, and given the need for constant imaging conditions between constructs to be able to compare the ratio values, the controls were infiltrated without p19 and at OD₆₀₀ 0.05, and imaged just before 48 hs post infiltration to avoid saturation in most cells, whereas pMMO proteins were infiltrated with p19 and at OD₆₀₀ 0.1 and a slightly longer time was allowed for more protein accumulation.

Table 2.8. Imaging conditions of roGFP2.

RoGFP2 was imaged using three tracks to measure the autofluorescence (Track 1), excitation at 405 nm (Track 2, predominant in the oxidized state), and excitation at 488 nm (Track 3, predominant in the reduced state).

	Track 1	Track 2	Track 3
Laser (nm)	405	405	488
Laser power (%)	0.8	1.2	0.6
Detection wavelength (nm)	410-470	500-555	500-555
Detector gain (V)	800	1000	1000

The images (single frames) were processed using a redox ratio analysis software (RRA) (Fricker, 2016) using the batch processing mode and loading the parameters from the roGFP2 database, following instructions in the manual. In brief, the software first processes the images applying a smoothing filter (an xy kernel of 3 was used), background subtraction (from a user selected area within the image), and autofluorescence subtraction. The autofluorescence bleed-through is calculated from an area of an image which has autofluorescence but no probe, and an autofluorescence correction value is applied to the 405 excitation channel (Track 2). The software then calculates a pixel-by-pixel ratio of the 405/488 signals, for which the recommended settings were used: an SD thresh of 2, which applies a threshold of two standard deviations above the background to mask low intensity pixels which could belong to structures

such as the vacuole, and a saturation of 0.9, which removes pixels with an intensity value above 90% of the full intensity scale. The software produces an average ratio value for each image and a pseudo-colour image showing the pixel-by-pixel ratio.

2.9 AnalyzER

The GFP-HDEL marker (OD₆₀₀ 0.05) was infiltrated into tobacco alone or in combination with RFP tagged pMMO proteins (OD₆₀₀ 0.1). Image acquisition was carried out as described in (Pain *et al.*, 2019) using the Zeiss 880 PlanApo 100x/1.46 NA oil immersion objective, an averaging of 2, dual channel acquisition using line switching (GFP detection using a band pass filter at 495-550 nm and RFP detection using a long pass filter at 570 nm) and Airyscan processing. 30-frame time series (825 ms frame time) were acquired using a constant zoom and maintaining imaging parameters throughout the experiment, and the green channel of each time series was imported into the AnalyzER software. The software was operated following the guidelines in the manual (Pain *et al.*, 2019) and using the default settings with the following changes: the pixel scaling and frame time were adjusted accordingly (0.09 $\mu\text{m}/\text{pixel}$ and 0.82 s), the full-width-half-minimum height (FWHM min), full-width-half-maximum height (FWHM max) and target width were set to 3, 7 and 5 respectively, the min A value for the classification of cisternae was set to 1 μm^2 , and the lag value for persistency analysis was set to 7 (equivalent to 5.7 seconds) (Figure 2.6). The resulting data was processed using the AnalyzER statistics package based on MATLAB (R2020b) to obtain boxplots of each variable, perform MANOVA, and carry out ANOVA with Tukey's HSD post-hoc comparisons.

For the tubule morphology analysis higher magnification images were used (pixel scaling 0.02 $\mu\text{m}/\text{pixel}$), and the files with both channels (GFP and RFP) were imported into AnalyzER. The template for skeleton extraction was chosen as the mean of both channels. Traces along a tubule were plotted using the "Trace", "selected" options in the software as explained in the manual.

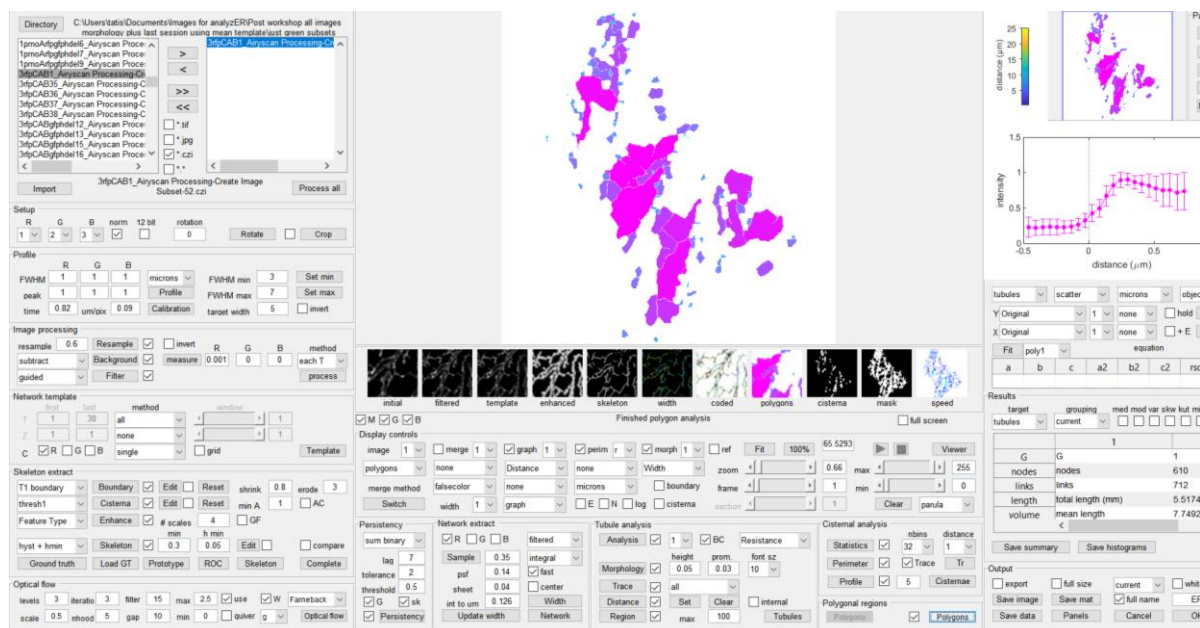


Figure 2.6. AnalyzER software user interface for ER structure and dynamics analysis.

Screenshot of the software user interface with the parameters used in the current study for analysis. Raw confocal data (.czi files) are loaded into the software, and the panels are run in a specific order to perform subsequently: image processing, skeleton extraction, optical flow analysis, persistency calculation, extraction of a weighted network, extraction of tubule metrics, cisternae analysis, and analysis of polygonal regions (Pain *et al.*, 2019). The central panel displays the image in the chosen processing step, with multiple options to alter the image display.

2.10 FRET-FLIM

Plants were infiltrated at Oxford Brookes University and transported to the Rutherford Appleton Laboratory at the Central Laser Facility (Harwell campus). Images were acquired as described in (Kriechbaumer *et al.*, 2016) by Charlotte Pain (Oxford Brookes), using a modified Nikon EC2 confocal scanning microscope which permitted two-photon fluorescence lifetime imaging (Botchway *et al.*, 2015). Excitation was produced with a 920 nm laser light generated by a titanium sapphire laser, emitting 200 fs pulses at 76 MHz, and focused with a water immersion Nikon VC/1.2 NA objective. The emitted light was collected without descanning, and passed through a BG390 (Comar) filter before reaching the detector. The detector was

linked to a time-correlated single-photon counting PC module (SPC830, Becker and Hickl) to generate the raw FRET-FLIM data.

The data was analysed by Verena Kriechbaumer (Oxford Brookes) using SPCImage 5.1 Software (Becker and Hickl), where lifetime exponential decay curves were traced for every pixel in the region of interest (in the nuclear envelope). The average excited-state lifetime values were obtained for each nuclear envelope.

2.11 RT-PCR

2.11.1 RT-PCR of transiently expressed RFP-polypeptide in tobacco

RNA was extracted from sections of infiltrated tobacco leaves at 2 days after infiltration (between 100-200 mg approximately), using the NEB Total RNA Miniprep Kit, following the manufacturer's instructions. Briefly, DNA/RNA protection agent was added to the samples and the plant material was lysed by mechanical disruption, following which RNA lysis buffer was added. The RNA was separated from most of the genomic DNA (gDNA) via a gDNA removal column and RNA was purified using 99% ethanol and an RNA purification column. Residual gDNA was removed by in column DNase I treatment. Following washes, the RNA was eluted in 50 µl nuclease-free water and the concentration was measured with a NanoDrop spectrophotometer. Additional gDNA removal was performed by in tube digestion with DNase I and samples were stored at -20 °C.

cDNA was synthesized using ProtoScript® II First Strand cDNA Synthesis Kit (NEB), according to the instructions in the kit, and using 1 µl of oligo(dT) primer and 1 µl of random primers. Parallel reactions were carried out without reverse transcriptase enzyme to control for genomic contamination. Primers were designed to target the ends of the synthesized polypeptide and previously designed primers by Charlotte Pain (Oxford Brookes) were used to amplify actin (Table 2.9). PCR was carried out using OneTaq® DNA Polymerase (NEB) and

bands were identified by gel electrophoresis (1% agarose gel) with 2 µl/100 ml SYBR Safe DNA Gel Stain (Invitrogen).

Table 2.9. Primers used for tobacco RT-PCR.

Forward and reverse primers used to amplify cDNA from leaf pieces transiently expressing RFP-polypeptide.

Primer name		Sequence (5'-3')
polyCshortF	Forward	ATTCGTTCGCCCCCTGAGTTC
polyCloverShortR	Reverse	ATTATACTCCAGCTTGTGGCCCA
Actinfor	Forward	TGGAAGTGAATGGTTAAGGCTGG
Actinrev	Reverse	TCTCCAGAGTCGAGCACAATACCG

2.11.2 RT-PCR of rice plants

RNA was extracted from leaf pieces of adult rice plants using the NEB Total RNA Miniprep Kit as described above, with the addition of previously freezing the tissue in liquid N₂ and performing some mechanical disruption of the tissue before adding the DNA/RNA protection reagent. This increased the final RNA yield compared to not including a previous mechanical disruption with liquid N₂. cDNA was performed using the ProtoScript® II First Strand cDNA Synthesis Kit (NEB) as described above, using specific primers for each pMMO subunit (Table 2.10).

Table 2.10. Primers used for rice RT-PCR.

Forward and reverse primers used for PCR following cDNA synthesis from rice leaf tissue expressing the polypeptide construct.

Primer name		Sequence (5'-3')
polyCf	Forward	ATGTCTGTCACAACTGAAACA
polyCr	Reverse	CGGCAACGCAATCCTTAC
polyAf	Forward	ATGTTCAACAAGCAAGAGTGG
polyAr	Reverse	GTATCACTTTGGTGGTGCTAT
polyBf	Forward	ATGGCTCGCCTTACCTC
polyBr	Reverse	ATGTCTCCAGCCACGAATTT

2.12 Substrates for propylene epoxidation assays

2.12.1 Preparation of plant lysates

Nicotiana benthamiana plants were infiltrated and leaf tissue was collected 2 days after infiltration (approximately 2 leaves, removing the midrib and veins). The tissue was ground to a fine powder with liquid N₂ using a mortar and pestle. The powder was homogenized with approximately 1.5 ml of non-denaturing lysis buffer (10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.5% (v/v) Nonidet P-40, 1 mM phenylmethylsulfonyl fluoride). For samples lysed with supplemental copper, Cu₂SO₄ solution was added to the lysis buffer with a final concentration of 40 µM. Samples were left on ice for 30 minutes and were centrifuged at 4000 rpm for 5 minutes. The supernatant was snap frozen and stored at -80°C.

2.12.2 Reduction of duroquinone

Duroquinol was prepared according to the protocol in (Zahn and Dispirito, 1996). In summary, 50 µl of HCl was added to 20 ml of ethanol (obtaining a final concentration of 3 mM HCl). 0.2g of duroquinone was added to the ethanol with stirring, following which 0.28g of sodium hydrosulfite was added to the solution with continuous stirring for 3 minutes. Subsequently, 0.07g of sodium borohydride was added with stirring, until the solution changed from a yellow colour to white. The solution was left stirring for 15 minutes and was then diluted with 150 ml H₂O. The solution was filtered through Whatman paper using suction filtration with the duroquinol remaining on the filter paper. The duroquinol was washed with 200 ml H₂O and dried with suction filtration for 15 minutes. The duroquinol was placed into an Eppendorf tube and stored at -80°C for later use (recommended to prepare fresh before use in assays).

3 Expression of pMMO proteins in the plant endoplasmic reticulum (ER)

3.1 Introduction

3.1.1 The role of the ER in protein synthesis and folding

The endoplasmic reticulum (ER) is a vast organelle present in eukaryotic cells that is involved in protein and lipid synthesis and export, protein folding, glycosylation, and more specialised functions in plants, such as lipid and protein body formation in seeds (Vitale and Pedrizzini, 2005; Schwarz and Blower, 2016). The ER has a complex and dynamic structure composed of two different types of interconnected elements: flat bilayered sheets or “cisternae” linked by tubules that also connect at three-way junctions to form a network.

The ER is the entry point to the secretory pathway. Proteins are targeted during translation to the ER membrane by the presence of an N-terminal signal peptide (SP). If they are destined for secretion, the signal peptide is cleaved and the protein is transported downstream to the Golgi, and subsequently to the plasma membrane and extracellular space (Schwarz and Blower, 2016). Proteins may also be targeted downstream from the ER to other organelles such as the peroxisome or the vacuole. Targeting to the ER occurs via the signal recognition particle (SRP), which binds the SP and links the nascent polypeptide to the ER membrane. Translocation into the ER then occurs through the Sec61 channel in an ancestrally conserved process highly similar to SRP targeting of proteins to the inner bacterial membrane (Rapoport, 2007). The ER is also the point of synthesis of most integral membrane proteins. Upon emergence of the first N-terminal transmembrane domain, nascent membrane proteins are recognised by the SRP and targeted in an analogous manner to soluble proteins. Following the synthesis of hydrophobic

transmembrane segments, these exit via a lateral gate in the Sec61 translocon and become embedded in the membrane (Rapoport, 2007).

Soluble proteins resident in the lumen of the ER, termed “reticuloplasmins”, are known to carry C-terminal retention signals such as KDEL and HDEL tetrapeptides that enhance their retrieval and retention in the ER (Denecke *et al.*, 1992; Napier *et al.*, 1992). The K/HDEL retention mechanism is based upon retrieval via the ERD2 receptor localized in the Golgi apparatus, where proteins are recycled in a retrograde manner (Silva-Alvim *et al.*, 2018). Similarly, some signals have been identified for ER retention of resident integral membrane proteins. For example, a cytosolic C-terminal dilysine motif (KKXX or KXKXX) is known to mediate retention of type I membrane proteins, which have their amino terminus in the lumen (Pagny *et al.*, 1999; Benghezal *et al.*, 2000), whereas a cytosolic diarginine motif causes retention of type II membrane proteins, which have their amino terminus in the cytosol (Pagny *et al.*, 1999). Another number of known and unknown protein sequences are thought to cause ER retention, and factors such as interaction with other resident ER proteins and oligomerization can also constitute a mechanism for retention (Pagny *et al.*, 1999; Boulafloous *et al.*, 2009).

In addition to its role in protein synthesis, the ER compartment houses several chaperones and plays a key part in protein folding and quality control. Some of the main chaperones include the luminal protein BiP from the Hsp70 family, protein disulfide isomerase (PDI), which aids in folding and in the formation of disulfide bonds as well as the reduction of non-native disulfide bonds, and the carbohydrate binding proteins calnexin and calreticulin, which bind glycoproteins in the membrane and in the lumen respectively (Braakman and Hebert, 2013). These and other chaperones help proteins to achieve their correct conformational shape and to avoid misfolding and aggregation, and form the first steps in the ER quality control pathway.

3.1.2 The plant ER as a compartment for foreign protein production

Plants are increasingly being used in research for the recombinant production of proteins for therapeutics or other biotechnological purposes, due to benefits such as low production cost,

high scalability, and ability to carry out post-translational modifications which are absent in prokaryotic systems (Shanmugaraj *et al.*, 2020). However, recombinant expression can often result in low protein yields due to gene silencing, or protein instability and degradation. Targeting proteins to the ER and retaining them via a HDEL or KDEL motif has in several cases resulted in increased protein accumulation levels (Vitale and Pedrizzini, 2005). As an example, single-chain variable fragment (scFv) antibodies were cloned in tobacco and expressed at high levels when retained in the ER lumen, yielding 4-6.8% of total soluble proteins, compared to 0.1% when cytosolic (Fiedler *et al.*, 1997; Conrad and Fiedler, 1998). It was hypothesized that the oxidizing environment of the ER and the presence of chaperones may aid in protein folding and stability, playing a role in the increased yields. Similarly, other studies have reported increased yields when producing proteins on the ER membrane. For example, the Nef protein (negative factor), a potential vaccine target for human immunodeficiency virus, was expressed in tobacco at low levels in the cytosol, whereas a threefold increase in protein accumulation was achieved via anchoring to the ER membrane by fusion to the ER integral membrane protein cytochrome b5 (Barbante *et al.*, 2008). It was thought that ER anchorage could have increased protein stability and conferred protection from the ubiquitin-proteasome system in the cytosol (Barbante *et al.*, 2008). In addition, the natural plasticity of the ER to environmental changes and to varying protein production levels also contributes to the organelle's capacity to harbor large amounts of foreign proteins (Vitale and Pedrizzini, 2005; Borgese *et al.*, 2006).

Another characteristic of the ER that renders it a favorable environment for protein bioengineering is its ability to promote the physical association of stable protein complexes or enhance the formation of metabolons. The ER surface has been reported to anchor metabolons through cytochrome P450 enzymes (Ralston and Yu, 2006), which are tethered to the ER membrane and interact both with membrane proteins such as their P450 reductase partners, and with cytosolic proteins. Examples of such metabolons include biosynthetic pathways for dhurrin and isoflavonoids (Nielsen *et al.*, 2007; Dastmalchi *et al.*, 2016). The ER membrane

also favours protein-protein interactions in the form of large multimeric protein complexes, such as the complex formed by ethylene receptor dimers and their interacting partners (Ju and Chang, 2012), and the endoplasmic reticulum membrane protein complex (EMC), an octameric chaperone which interacts also with neighbouring proteins to aid translocation and insertion of proteins on the ER membrane (Bai *et al.*, 2020).

3.1.3 Methods for testing protein-protein interactions

Determination of protein-protein interactions (PPIs) is necessary for the identification of metabolons and protein complexes such as those present on the ER membrane, and it is relevant in this study as pMMO is a multimeric membrane bound complex. Examples of common methods used to determine protein-protein interactions *in vivo* include the yeast two-hybrid method (Y2H), bimolecular fluorescence complementation (BiFC), and Förster resonance energy transfer (FRET) (Xing *et al.*, 2016). The yeast two-hybrid assay makes use of a transcription factor (GAL4) which contains two domains necessary for its function: a DNA binding domain and an activation domain. The domains are synthetically separated and fused to each protein of interest. In the event of an interaction between the target proteins, the domains of GAL4 become close enough to conform a functional version of GAL4 and activate the transcription of a reporter gene. The BiFC method follows a similar principle using a GFP molecule split into two fragments. By fusing the proteins of interest to either fragment, an interaction can be detected as this causes the fragments to be in close proximity and therefore reassemble into a functional GFP (Xing *et al.*, 2016).

In this study FRET-FLIM (Förster resonance energy transfer by fluorescence lifetime imaging) was used to test PPIs, due to the high sensitivity of the method and the availability of fluorescent protein fusions which were created for localization studies. FRET-FLIM is based on the occurrence of FRET, which is the non-radiative transfer of energy between two fluorophores found in close proximity. For FRET to occur the fluorophores must have significant spectral overlap, that is the emission spectrum of the “donor” fluorophore must

overlap with the absorption spectrum of the “acceptor” fluorophore for energy to be transferred from the donor to the acceptor (Figure 3.1) (Wallrabe and Periasamy, 2005).

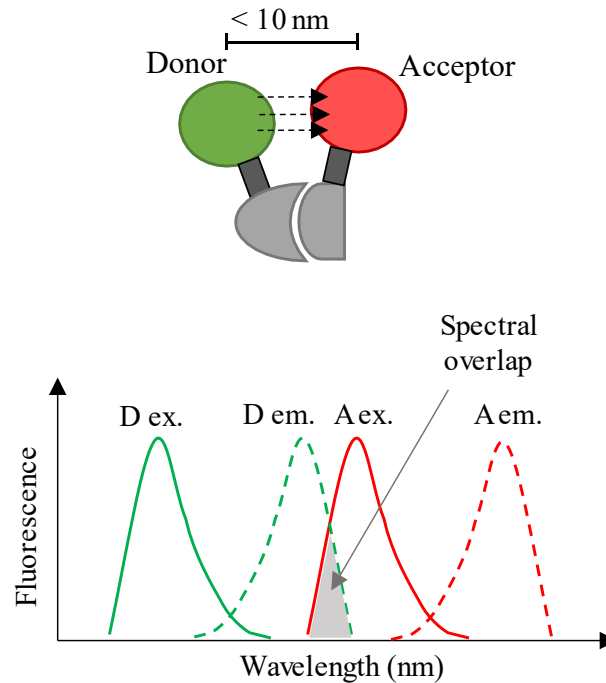


Figure 3.1. The principles of FRET.

Target proteins (shown in grey) are fused to donor and acceptor fluorophores. In the case of the fluorophores being at a distance closer than 10 nm driven by an interaction between the target proteins, non-radiative transfer of energy occurs from the donor to the acceptor. The emission spectrum of the donor (D em.) must sufficiently overlap with the excitation spectrum of the acceptor (A ex.) in order for FRET to occur. Adapted from (Broussard and Green, 2017).

FRET efficiency (E) is inversely proportional to the sixth power of the distance between fluorophores (R), as shown in:

$$E = \frac{1}{1 + (R/R_0)^6}$$

Due to the strong dependance on distance, FRET is only detectable when the proteins are found at <10 nm from each other. FRET efficiency is also dependent on the quantum yield of the donor, and the alignment of the transition moments of the fluorophores, which are included in the Förster distance parameter (R_0) (Corry *et al.*, 2006). The transfer of energy from the donor to the acceptor causes quenching of the donor fluorescence, which is measured in intensity-

based (or spectral) FRET. In a biological system, if fusion of donor and acceptor fluorophores to different subunits of interest renders them close enough for FRET to occur, it is likely that there is an interaction between the proteins of interest.

FRET-FLIM measures the decrease in lifetime (average time that a fluorescent species remains in the excited state) of the donor molecule which occurs as a result of FRET (Schoberer and Botchway, 2014). The lifetime (τ) of a fluorophore will determine the fluorescence intensity decay over time, which is governed by the exponential function:

$$I(t) = I_0 e^{\frac{-t}{\tau}},$$

where I_0 is the initial fluorescence intensity, t is the time, and $I(t)$ is the fluorescence intensity over time (Datta *et al.*, 2020). FRET-FLIM is a more technically challenging method than spectral FRET, however, its advantages include a higher sensitivity and independence on the relative protein concentrations of donor and acceptor molecules (Xing *et al.*, 2016), which tend to be very variable in transient co-transformation experiments in plants.

3.1.4 Methods for determination of membrane topology

Protein-protein interactions of membrane proteins are highly dependent on membrane topology. Integral membrane proteins span the membrane one or more times according to the number of hydrophobic transmembrane segments present in the protein sequence. Membrane topology shows the number of times the protein crosses the bilayer and the position of each fragment with respect to the cytosol or lumen of the organelle. Determining membrane topology can be important to understand the subcellular environment of the different protein domains as well as the regions that might interact with other resident proteins.

In this study a redox-based topology analysis to determine the membrane topology of the recombinant pMMO proteins was applied. This method makes use of the difference in redox potential of glutathione (E_{GSH}) between the cytosol and the lumen of the ER: the ER lumen constitutes an oxidizing environment with a similar proportion of reduced glutathione (GSH)

to oxidized glutathione (GSSG), whereas the cytosol is a reducing environment with a much higher proportion of GSH (Brach *et al.*, 2009). The method uses a redox-sensitive GFP molecule (roGFP2), which contains two mutations to create surface exposed cysteine residues, which become oxidized or reduced according to the surrounding redox environment (Hanson *et al.*, 2004). The cross-links between cysteine residues governed by the oxidation state slightly modify the protein conformation near the chromophore and confer roGFP a modified excitation spectrum. Specifically, roGFP2 fluoresces at a higher efficiency with excitation at 405 nm in an oxidized environment, and at a higher efficiency with excitation at 488 nm in a reduced environment, therefore the 405/488 nm ratio provides an indication of the oxidation state of the protein (Brach *et al.*, 2009).

3.1.5 Objectives

The aim of this chapter was to recombinantly express all three bacterial pMMO subunits in the ER membrane of tobacco plants using agrobacterium-mediated transient transformation of tobacco to express the subunits from the nuclear genome. The ER endomembrane compartment provides a large surface area which would potentially accommodate high amounts of the recombinant proteins, resembling the intracytoplasmic membranes present in methanotroph bacteria which house the native pMMO complex. The ER also comprises a suitable environment to aid protein folding and to promote protein-protein interactions, which poses favourable conditions for pMMO complex assembly.

The specific objectives were:

- 1) To express fluorescently tagged pmoC, pmoA and pmoB from separate promoters in tobacco plants, using transient transformation.
- 2) To assess subcellular localization of the subunits using confocal microscopy and to target all proteins to the ER with appropriate signal peptides and retention signals.
- 3) To test protein-protein interactions and membrane topology via FRET-FLIM.

- 4) To confirm membrane topology with the use of redox-sensitive fluorescent proteins (roGFP2).
- 5) To test transient co-expression of all three proteins in individual cells using three colour imaging.

3.2 Results

3.2.1 Subcellular localization of pMMO proteins in tobacco cells

3.2.1.1 Protein structure and prediction of localization in Eukarya

Determining the subcellular localization of recombinant proteins is a key step when creating a bioengineered system. The localization within the cell can impact upon the presence of cofactors, pH levels, interactions with other proteins, and in this case correct complex assembly. Information derived from primary sequence data, as well as structural data and bioinformatics tools were used to predict the subcellular localization of the foreign pMMO proteins in tobacco.

The bacterial pMMO complex from *Methylosinus trichosporium* OB3b has been crystalized previously (PDB ID 3CHX), revealing the structure of the pmoC, pmoA and pmoB subunits (Hakemian *et al.*, 2008), listed hereafter according to the gene order in the bacterial operon (pCAB). PmoC and pmoA contain transmembrane domains throughout the length of the proteins (Figure 3.2 A). However, the final 80 amino acids of pmoC are unmodelled in the crystal structure. Indeed, all crystal structures of pMMO obtained to date contain unmodelled regions of pmoC of varying sizes due to protein disorder (Koo and Rosenzweig, 2021). However, computational algorithms consistently predict two additional transmembrane domains in pmoC from 3CHX, yielding a total of six transmembrane domains (Figure 3.2 B and Supplementary figure 3.1 A).

The pmoA subunit contains several transmembrane domains at inclined angles with respect to each other. Following the fifth transmembrane domain and a beta hairpin, the crystal structure shows a loop that traverses the membrane (Figure 3.2 A). Computational predictions vary in the number of transmembrane domains assigned to pmoA (Supplementary figure 3.1 B). To coincide with the topology adopted in the crystal structure, the loop is here considered as a sixth transmembrane domain, yielding 7 transmembrane domains in pmoA (Figure 3.2 B).

In contrast to the subunits formerly described, pmoB contains two large periplasmic domains forming beta barrels separated by two transmembrane domains (Figure 3.2). A signal peptide is located at the N-terminus comprising residues 1-39 which directs the protein to the periplasm and is cleaved to activate the enzyme (Lieberman and Rosenzweig, 2005). Two copper coordinating sites are present in the enzyme located in the pmoC and pmoB subunits.

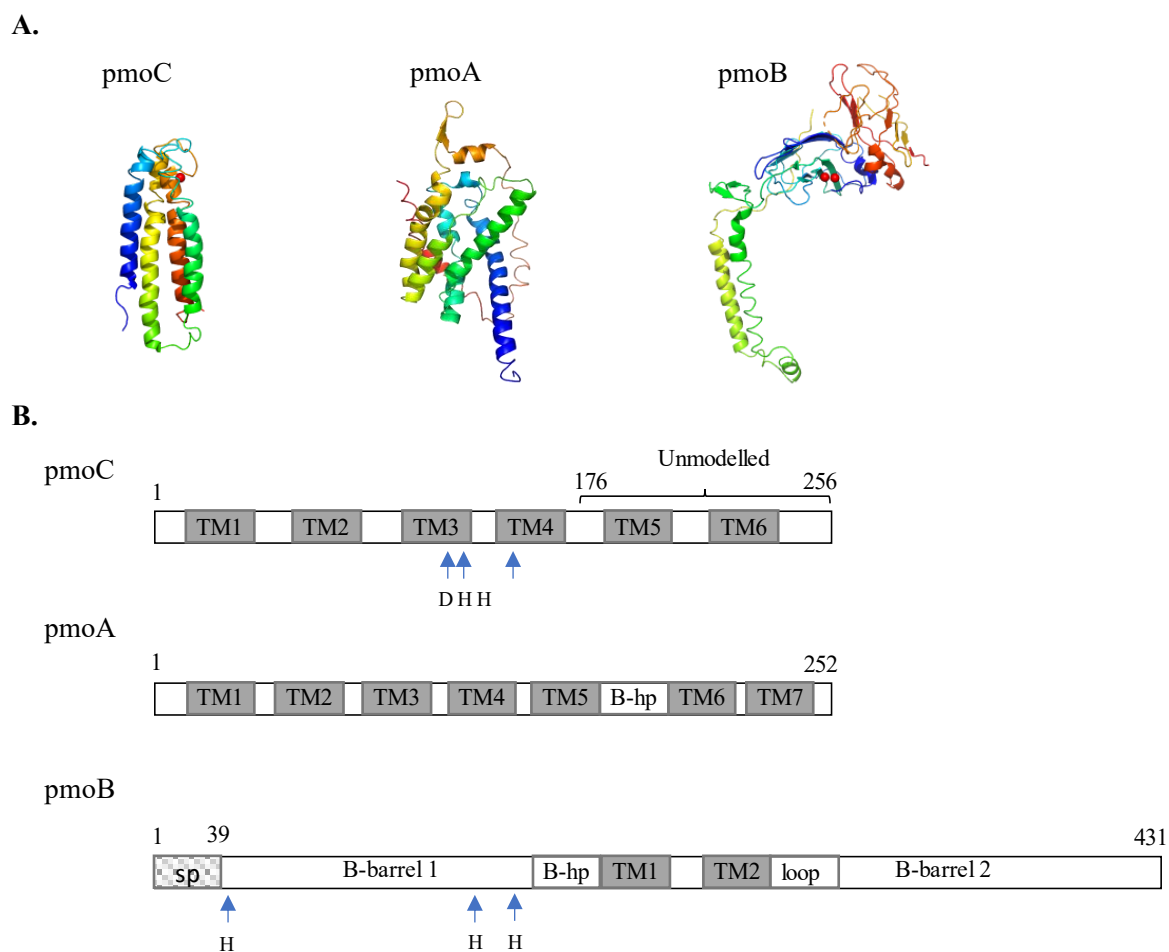


Figure 3.2. Structure of pmoC, pmoA and pmoB proteins.

(A) Cartoon representation of pmoC, pmoA and pmoB from *Methylosinus trichosporium* OB3b (PDB 3CHX) highlighting secondary structure elements. (B) Schematic diagram of the pMMO proteins and their structural elements. The unmodelled region of pmoC is predicted to contain two additional transmembrane domains using the bioinformatic algorithm CCTOP. Blue arrows indicate residues participating in copper binding sites: aspartic acid (D 129), and histidine (H 133 and H 146) form the copper site at pmoC (CuC); three histidine residues (H 40, H 144, H 146) compose the copper site at pmoB (CuB). TM = transmembrane domain, B-hp = beta hairpin, sp= signal peptide.

The bioinformatics algorithm DeepLoc 1.0 (Almagro Armenteros *et al.*, 2017) was used to predict the subcellular localization of pMMO proteins in eukaryotic cells. The algorithm performs a domain-specific prediction based on primary sequence information and using deep neural networks. The prediction yielded a likelihood of almost 1 of all three proteins being folded as membrane proteins in Eukarya (Table 3.1 A), and a high likelihood for the localization of pmoC and pmoA in the ER, and for the localization of pmoB in both the mitochondria and in the ER (Table 3.1 B).

Further analysis on the pmoB signal peptide was carried out using SignalP-5.0 (Almagro Armenteros *et al.*, 2019), which can predict the presence of a signal peptide across wide phylogenetic diversity (Archaea, Gram-positive bacteria, Gram-negative bacteria and Eukarya). The algorithm correctly predicted the presence of a signal peptide in Gram-negative bacteria, with a cleavage site between residues 39 and 40, indicated by a peak in the green dotted line (Figure 3.2 B). Interestingly, the same sequence was not predicted to act as a signal peptide in Eukarya (Figure 3.2 A). Some probability for a signal peptide was present (red line), however there was no likelihood for a cleavage site (Figure 3.2 C). It is thus possible that the bacterial signal peptide will not direct pmoB to the secretory pathway in plant cells.

Table 3.1 Subcellular localization predictions of pMMO proteins in Eukarya.

(A) Likelihood of folding as a membrane protein in Eukarya using DeepLoc 1.0, (B) Subcellular localization prediction of pMMO proteins in Eukarya using DeepLoc 1.0 showing the likelihood calculated for each of 10 subcellular compartments.

A.

	pmoC	pmoA	pmoB
Soluble	0	0.0001	0.0115
Membrane	1	0.9999	0.9885

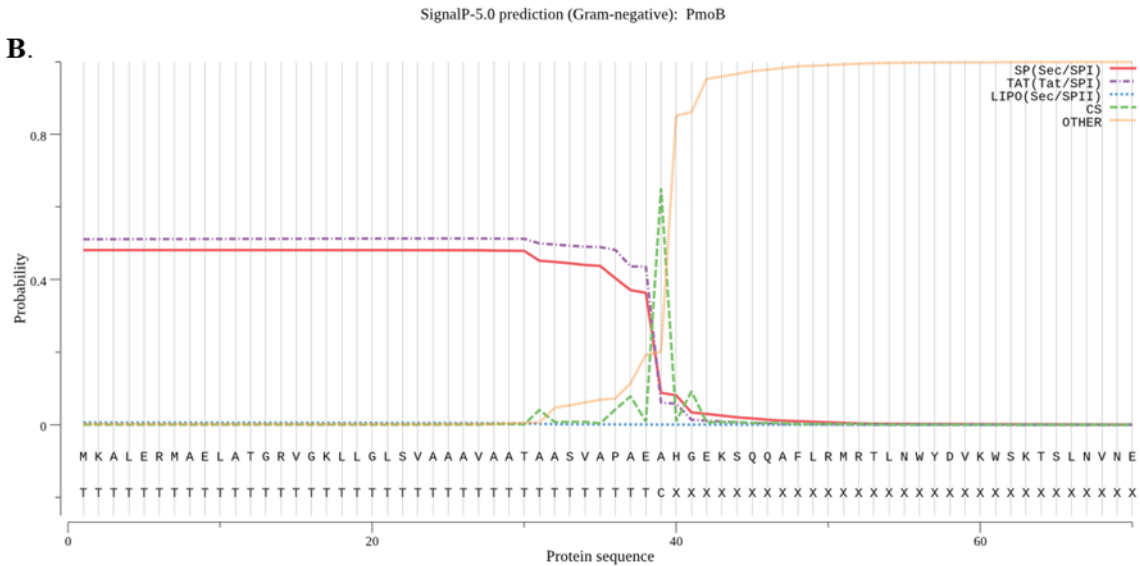
B.

Localization	pmoC	pmoA	pmoB
ER	0.8633	0.4197	0.4181
Plastid	0.0564	0.0008	0.0188
Mitochondrion	0.046	0.0017	0.4478
Cell membrane	0.0129	0.1074	0.0617
Golgi apparatus	0.0118	0.2307	0.0032
Lysosome/Vacuole	0.0081	0.2393	0.039
Peroxisome	0.0013	0.0002	0.0067
Nucleus	0.0001	0	0.0002
Cytoplasm	0.0001	0	0.0004
Extracellular	0	0	0.0042

A.

	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Gram-ve	0.4805	0.5132	0.0052	0.001
Eukarya	0.3513	-	-	0.6487

B.



C.

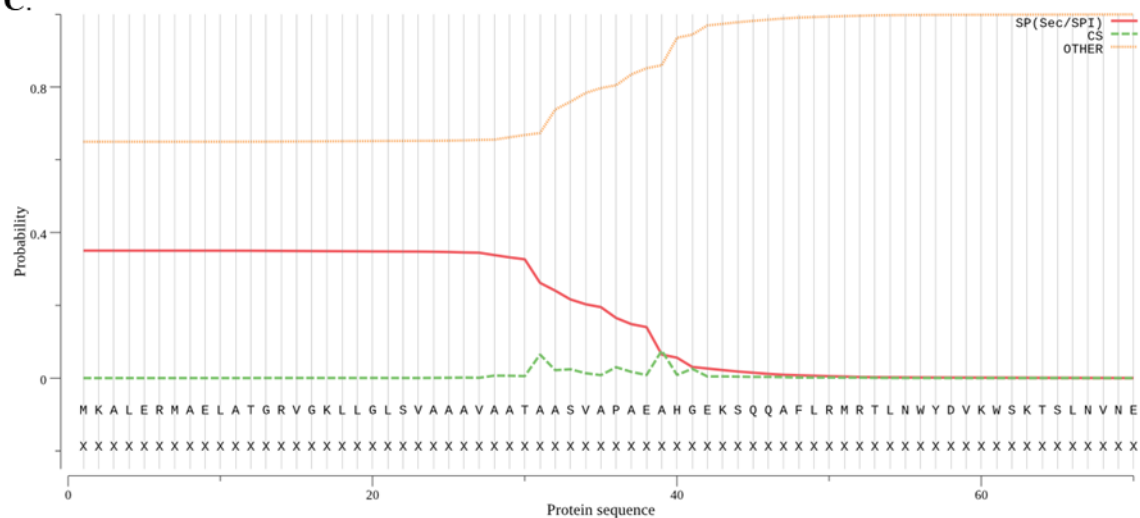


Figure 3.3. Prediction of pmoB signal peptide in Bacteria and Eukarya with SignalP-5.0.

(A) Likelihood of a signal peptide of the pmoB protein in Gram-negative bacteria and in Eukarya, final classifications are highlighted in bold. (B) Prediction of a signal peptide in Gram-negative bacteria. (C) Prediction of a signal peptide in Eukarya. In (B) and (C) the likelihood of a signal peptide is represented in red, the likelihood of a cleavage site in green, and the likelihood of signal peptide absence in yellow.

3.2.1.2 PmoC and pmoA localize to the plant ER

Predictions based on the protein sequences suggested that pmoC and pmoA should localize to the ER. To test the subcellular localization experimentally, the subunits were cloned into binary plant expression vectors under a strong constitutive promoter (35S), with a GFP tag on either the N- or C-terminus. The expected localization of the recombinant fusion proteins on the ER membrane, assuming that membrane insertion and folding occurs analogously as in the bacteria, is shown in Figure 3.4, where pmoC contains six transmembrane domains with both termini facing the cytosol, and pmoA contains seven transmembrane domains with the N-terminus facing the cytosol and the C-terminus facing the ER lumen.

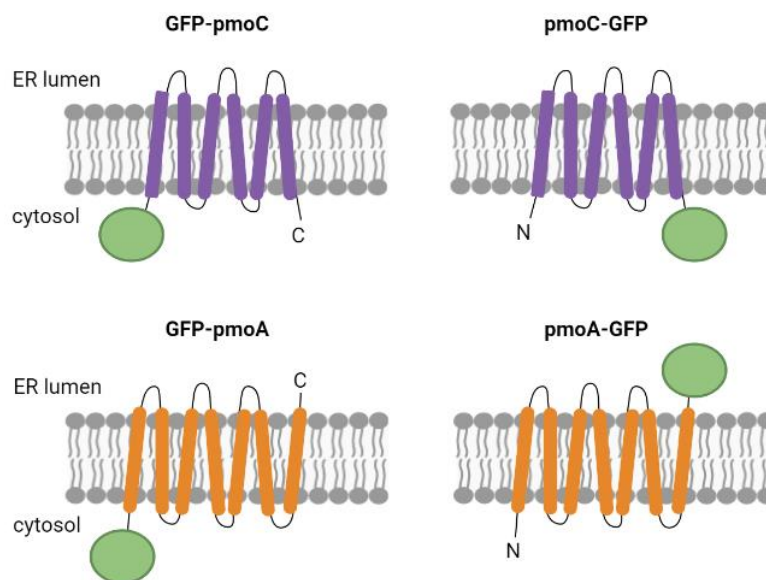


Figure 3.4. Expected folding of pmoC and pmoA fusion proteins to GFP on the ER membrane.

Diagram of the expected topology of pmoC and pmoA fusion proteins with GFP on the ER membrane. The position of GFP (green circle, not to scale) on either the cytosol or the ER lumen is shown.

Localization was assessed by transient transformation of tobacco and imaging at two days after infiltration, using a Zeiss 880 confocal laser scanning microscope with an Airyscan detector for improved resolution. Both GFP-pmoA and pmoA-GFP showed high expression levels in tobacco cells and localization at the endoplasmic reticulum, in accordance with the

computational predictions. The ER labelling is characterized by a well-defined reticulate network, composed of sub-regions termed tubules and cisternae (Figure 3.5). Localization at the endoplasmic reticulum was confirmed by co-infiltration with the ER luminal marker RFP-HDEL. Co-expression resulted in an overlap between the green and red signals, highlighted in a high magnification section of the network in the merged channel (Figure 3.6). Moreover, line profiles showed coinciding peaks and valleys of intensity of both channels across a linear transect.

GFP-pmoC and pmoC-GFP also localized to the ER, as could be seen by co-localization with RFP-HDEL (Figure 3.7). Line profiles also showed coincident increases and decreases in intensity between the channels.

All constructs were infiltrated using an optical density of agrobacterium culture (OD_{600}) of 0.05, given that at OD_{600} 0.1 more damage to the plant cells was observed. At OD_{600} 0.05, most cells expressing pmoA did not show any disruption to the ER by visual inspection, suggesting that plant cells are able to express pmoA without apparent damage to subcellular morphology. Approximately 50% of cells expressing GFP-pmoC presented some level of disruption to the ER. In the case of pmoC-GFP nearly all cells showed disruption of the ER network (Figure 3.7). The observations of ER damage at the subcellular level reflected damage to the leaf tissue seen at the macroscopic level. Severe tissue necrosis could be observed in leaf sections at two days after infiltration with GFP-pmoC at both OD_{600} 0.05 and 0.1, whereas some tissue necrosis was observed with GFP-pmoA at OD_{600} 0.1 and no visible damage at OD_{600} 0.05 (Figure 3.8), indicating that pmoC is somewhat more damaging to the plant. The ER membrane marker GFP-Calnexin (GFP-CNX) was infiltrated as a control, showing no visible tissue damage.

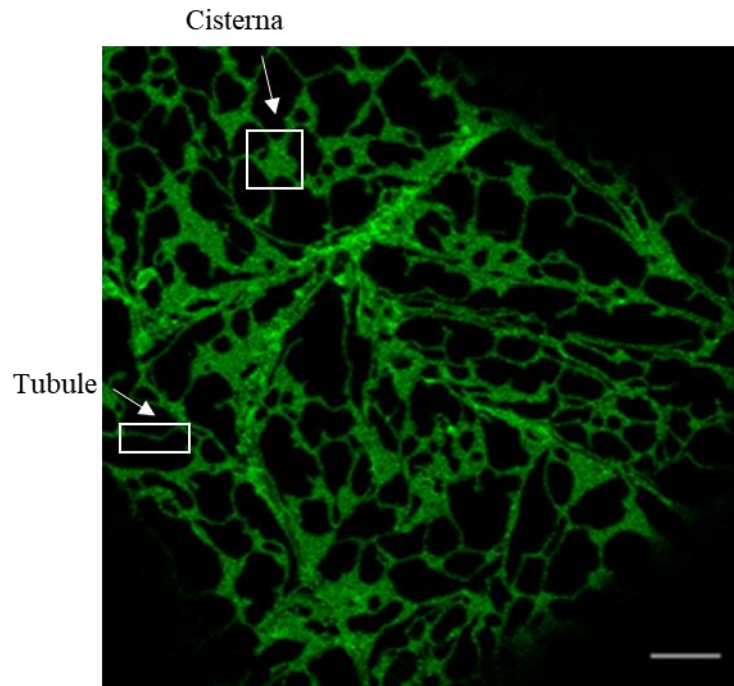


Figure 3.5. Expression of pmoA in the plant ER.

Transient expression of GFP-pmoA in tobacco epidermal leaf cells imaged by confocal microscopy with Airyscan. ER localization is seen through labelling of its defined reticulate structure composed of two major sub-regions: tubules and cisternae. Scale bar = 5 μm .

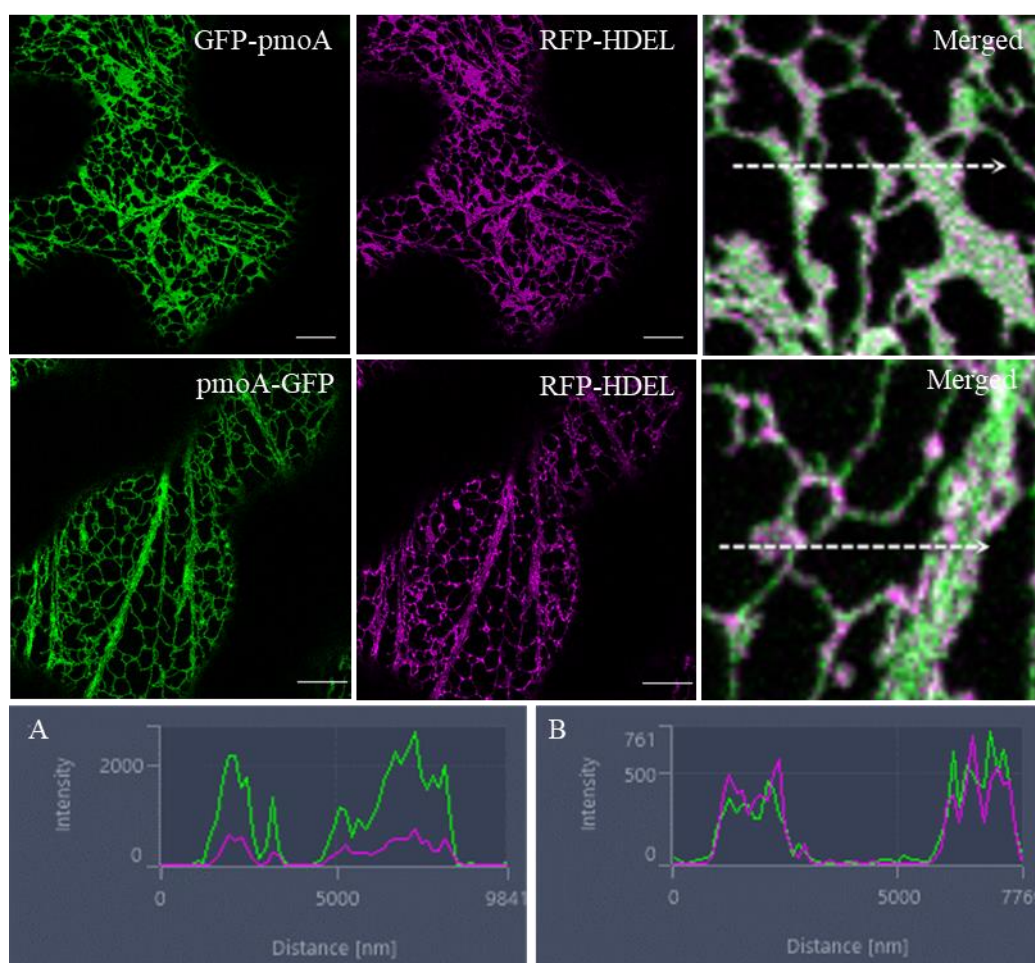


Figure 3.6. Transient expression of pmoA with RFP-HDEL in tobacco cells.

Co-localization of pmoA with the ER marker RFP-HDEL in tobacco cells, at 2 days after infiltration, showing an overlap between the GFP and RFP signals in both N- and C-terminal GFP tags. The merged images show the combined channels in a magnified area of the cell. Arrowheads indicate the transect (around 10 μm) for which a line profile is shown. (A) Line profile of intensities of GFP-pmoA and RFP-HDEL, (B) line profile of intensities of pmoA-GFP and RFP-HDEL, both with coinciding peaks in intensity. Scale bar = 10 μm .

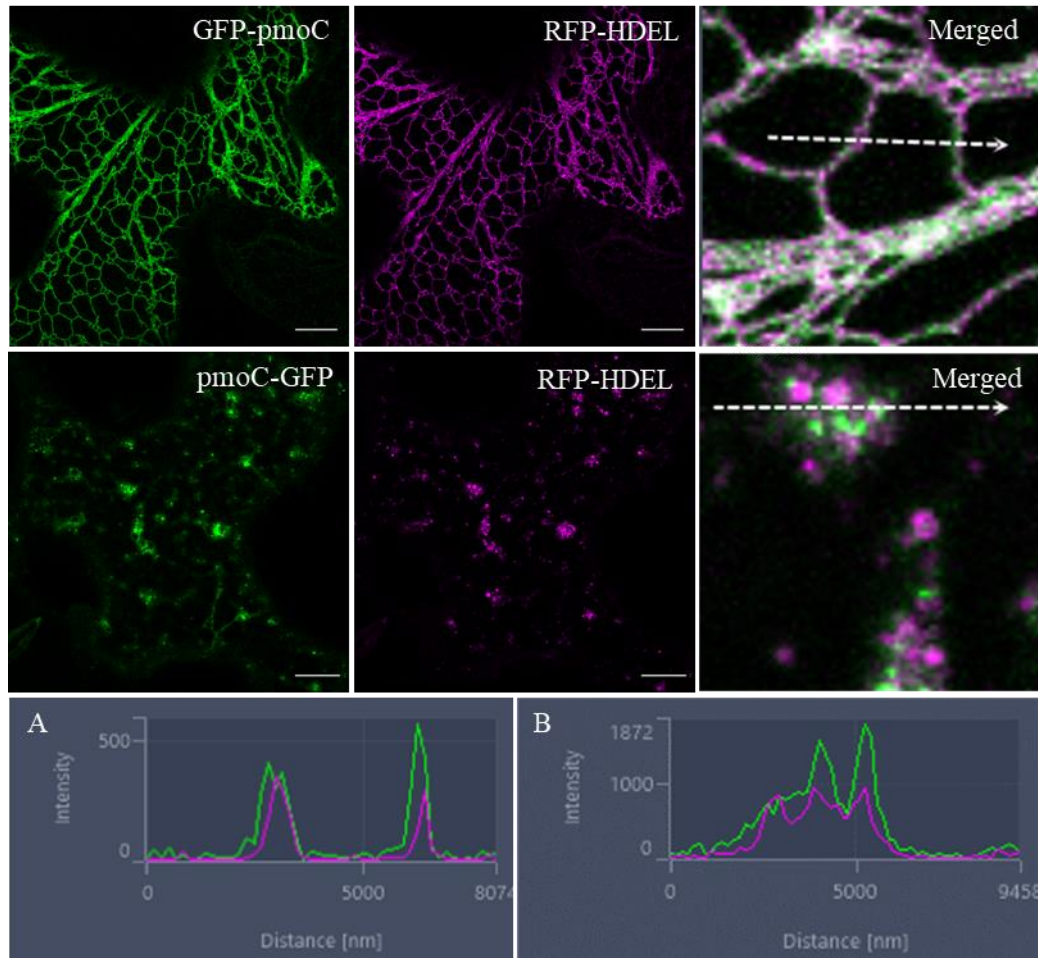


Figure 3.7. Transient expression of pmoC with RFP-HDEL in tobacco cells.

Co-expression of GFP-pmoC and pmoC-GFP with the ER marker RFP-HDEL in tobacco leaf epidermal cells, at 2 days after infiltration. Both N- and C-terminally tagged constructs show co-localization with RFP-HDEL, with GFP-pmoC presenting a normal ER morphology and pmoC-GFP inducing disruption of the ER network. Merged images show the combined channels in a region of the cell with high magnification. Arrowheads show the position of the line profiles. Line profiles of GFP-pmoC and RFP-HDEL (A) and pmoC-GFP and RFP-HDEL (B) reflect the signal overlap observed in the merged images. Scale bar = 10 μm .

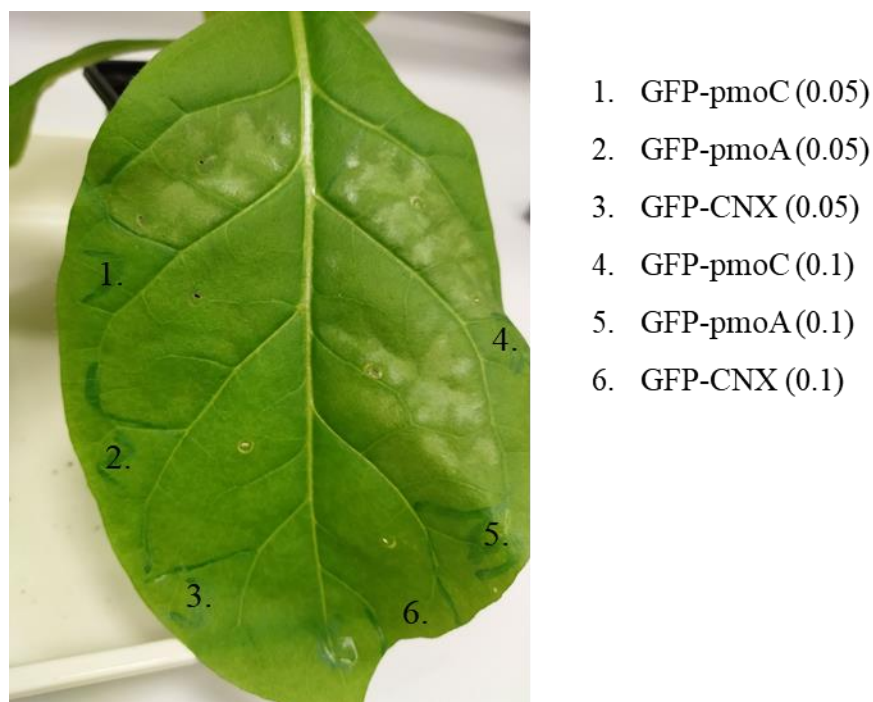


Figure 3.8. Tissue necrosis of infiltrated leaf with pmoC and pmoA.

GFP-pmoC, GFP-pmoA and the ER membrane marker GFP-CNXX were infiltrated at OD₆₀₀ 0.05 and OD₆₀₀ 0.1 in different leaf sections (1-6). Necrosis can be observed as severe thinning of the leaf tissue for GFP-pmoC at both ODs (1,4), and for GFP-pmoA at the higher OD (5) at 2 days after infiltration.

3.2.1.3 PmoB localizes to the cytosol and to bright puncta in a signal peptide-dependent manner

Computational predictions suggested the localization of pmoB as likely to be in the ER or in the mitochondria in eukaryotic cells. To test the localization of pmoB, the subunit was cloned into a plant expression vector under a strong constitutive promoter, with either an N- or C-terminal GFP tag. Following the topology dictated by the crystal structure, both pmoB termini are expected to be localized in the ER lumen or the mitochondrial intermembrane space (Figure 3.9).

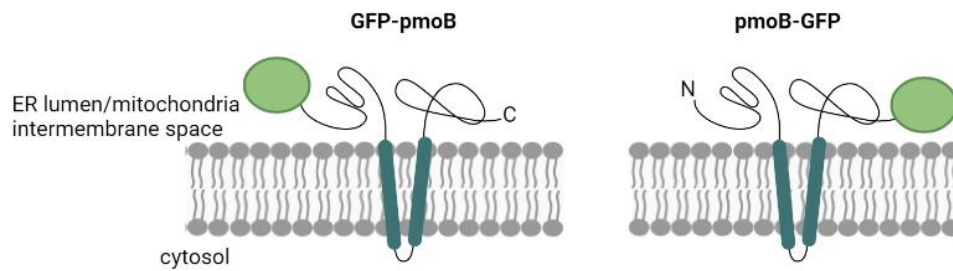


Figure 3.9. Expected orientation of pmoB fusion proteins to GFP on the ER or mitochondria membrane.

The expected structure and localization of pmoB fused to GFP (green circle, not to scale) on the ER or mitochondria, with both N- and C- termini facing either the ER lumen or the mitochondria intermembrane space.

Transient expression of GFP-pmoB in tobacco showed a diffuse pattern and negative staining of cellular vesicles, which is characteristic of cytosolic localization (Figure 3.10-top row). Co-localization between GFP-pmoB and RFP-HDEL yielded a partial overlap in the merged signal as the ER and cytosolic signal cannot be spatially distinguished with the resolution of the confocal system even with the Airyscan detector. The typical width of the ER tubule (40 nm (Pain *et al.*, 2019)) means the cross-section of an ER tubule is completely contained within the theoretical point spread function (PSF) of the microscope, which will therefore also include the surrounding cytosolic signal. In confirmation of this, the line profile of intensities across a linear section of the cell shows distinct peaks and valleys of the RFP signal on the ER network, and a relatively constant intensity of the green signal whose pattern does not coincide with the changes in red (Figure 3.10 A).

On the other hand, transient expression of pmoB-GFP revealed a combination of cytosolic signal and discrete brighter punctate objects (Figure 3.10 bottom row). A line profile encompassing some puncta shows a negative correlation between the objects (green) and the ER signal (magenta) (Figure 3.10 B), suggesting that these objects do not reside in the ER.

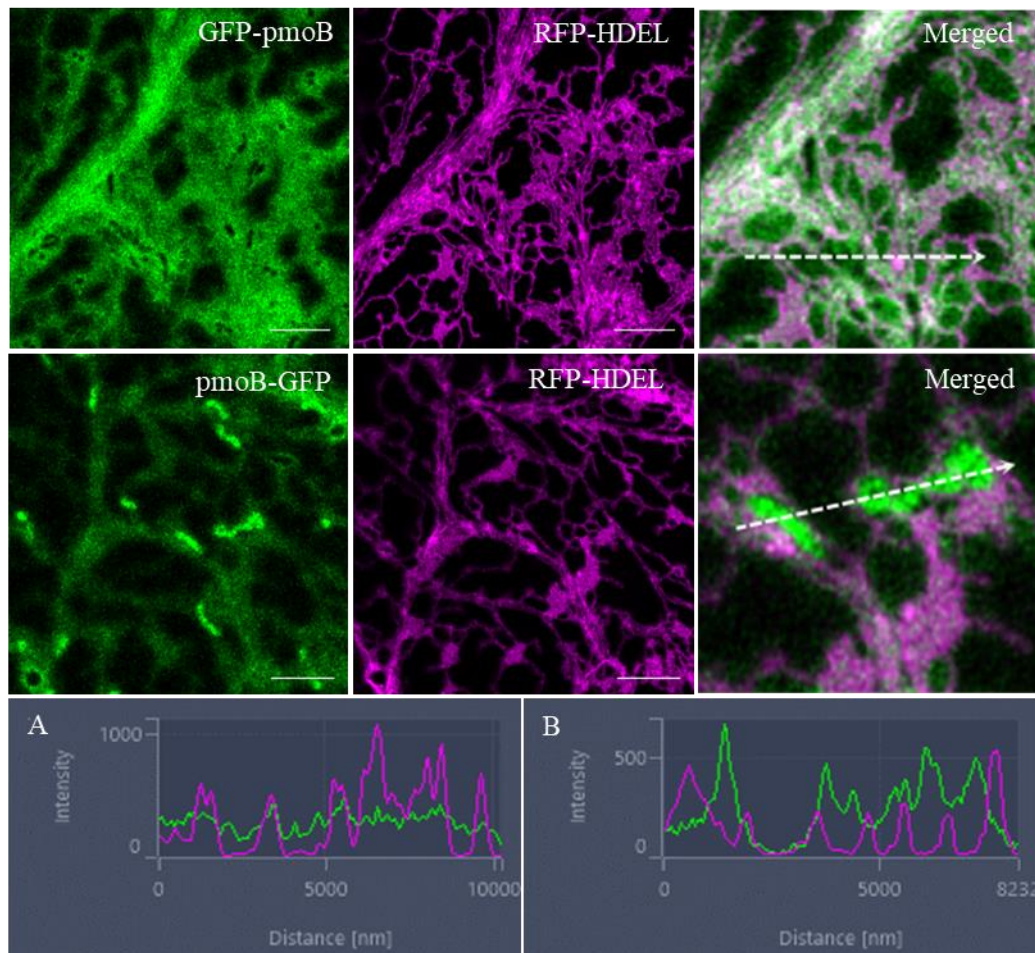


Figure 3.10. Transient expression of pmoB with RFP-HDEL in tobacco.

GFP-pmoB shows cytosolic localization with a partial overlap with the RFP-HDEL signal (merged channel and line profile (A)). PmoB-GFP localizes to the cytosol and to bright puncta, whose intensity negatively correlates with RFP-HDEL intensity peaks (B). Scale bar = 5 μ m.

Given the role of signal peptides in targeting, it was hypothesized that the pmoB signal peptide may be contributing to the formation of punctate objects, and that its function is masked when a fluorophore is placed N-terminally resulting in the completely cytosolic localization seen with GFP-pmoB. To determine experimentally whether this peptide influences the localization in plants, a truncated version of pmoB (“Btrunc1”) was created, removing the full-length signal peptide, creating a protein which starts on the His40 residue (Table 3.2). Three other truncated versions of pmoB were designed: one includes 7 amino acids N-terminal to His40 (“Btrunc2”), another one leaves a stretch of hydrophobic amino acids (mainly alanine and valine), which

could play a part in aggregation (“Btrunc3”), and a final truncation removes additional residues upstream of His40 (“Btrunc0”) (Table 3.2).

Following transient transformation of tobacco, all constructs localized to the cytoplasm and no formation of bright punctate objects was observed (Figure 3.11). This suggests that the truncated amino acids are necessary for the formation of punctate objects, whether this is the full signal peptide or solely residues from the first 19 amino acids truncated in Btrunc3. Interestingly, expression levels between the constructs differed, with a much stronger expression in Btrunc0 with respect to the rest, which could be related to protein stability.

Table 3.2. Truncations of the pmoB signal peptide.

Truncated versions of pmoB showing amino acids that were removed in dark grey. The first amino acid of the mature protein sequence in bacteria (His40) is shown in red. A non-native methionine residue was added for translation initiation and is shown in blue. All proteins localised to the cytoplasm but with varying expression levels.

Construct	Truncated amino acids (dark grey)	Expression
Btrunc0-GFP	MKALERMAELATGRVGKLLGLSVAAA VAATAASVAPAEA HGEKSQQMAFLR	High
Btrunc1-GFP	MKALERMAELATGRVGKLLGLSVAAA VAATAASVAPAEA MHGEKSQQAFRLR	Low
Btrunc2-GFP	MKALERMAELATGRVGKLLGLSVAAA VAATAASVAPAEA MHGEKSQQAFRLR	Medium
Btrunc3-GFP	MKALERMAELATGRVGKLLGLSVAAA VAATAASVAPAEA MHGEKSQQAFRLR	Very low

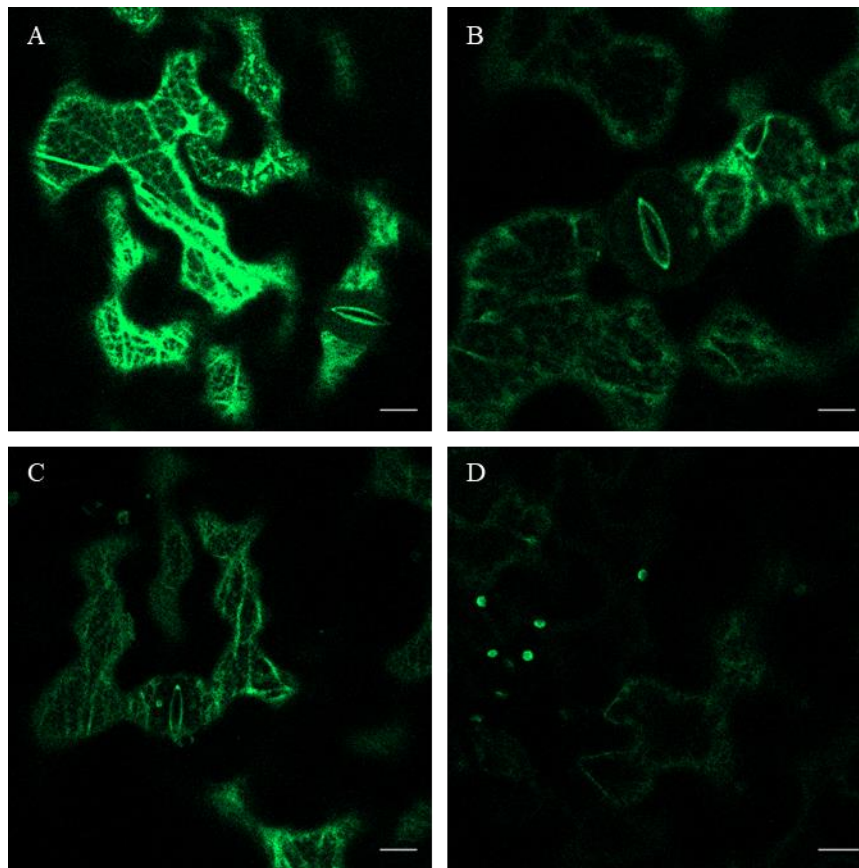


Figure 3.11. Transient expression of Btrunc-GFP constructs in tobacco.

Expression of Btrunc0-GFP (A), Btrunc1-GFP (B), Btrunc2-GFP (C), and Btrunc3-GFP (D). All show cytosolic localization with varying levels of expression. Scale bar = 20 μ m.

3.2.1.4 PmoB puncta form ring structures which co-localize with mitochondria

Higher magnification imaging with Airyscan detection revealed that pmoB labelled objects form mobile ring structures (Figure 3.12 A). Co-localization of pmoB-GFP with different organellar markers was performed to determine if the protein is being targeted to a specific subcellular compartment. Airyscan imaging shows that pmoB-GFP does not co-localize with the ER marker RFP-HDEL or the Golgi marker ST-RFP (Figure 3.12 B and C). Instead, co-localization is observed with the mitochondria marker RFP-ATPase (Figure 3.12 F and I). Analysis of time series also supports the co-localization of pmoB-GFP and RFP-ATPase as both proteins move together, as exemplified in Figure 3.12 J-L.

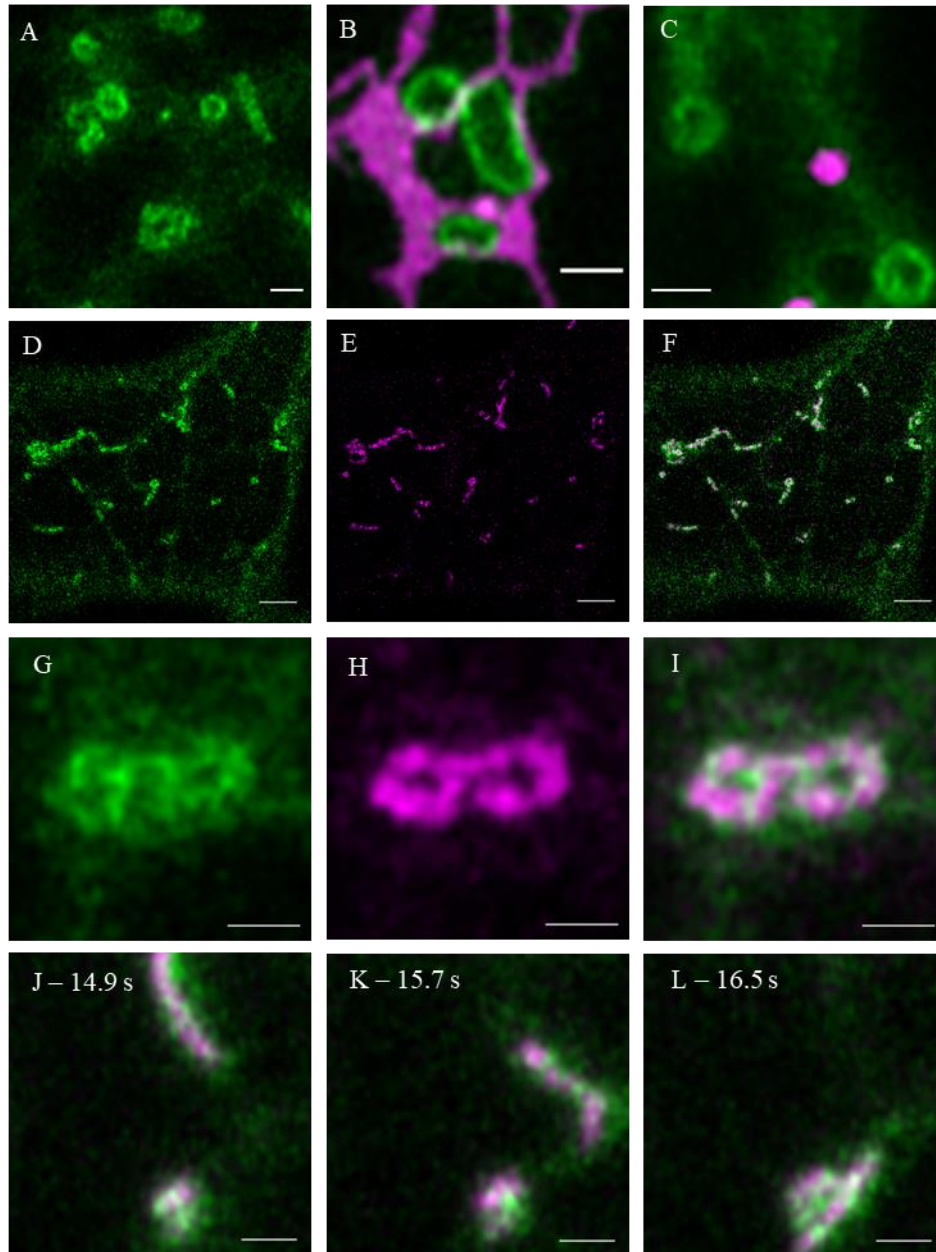


Figure 3.12. Co-expression of pmoB-GFP with organellar markers.

A) PmoB-GFP forms diverse ring and elongated structures. Co-expression and high magnification imaging of pmoB-GFP with RFP-HDEL (B) and ST-RFP (C). Co-expression of pmoB-GFP with RFP-ATPase, showing the green, red and merged channels (D-F). Co-expression and high magnification imaging of pmoB-GFP with RFP-ATPase, showing the green, red and merged channels (G-I), and showing the merged channel for different time points in a time series (J-L). Scale bar = 5 μm (A, D, E, F) and 1 μm (B, C, G-L).

3.2.2 PmoB does not relocate to the ER when co-expressed with pmoA

The crystal structure of the pMMO complex shows a tight association between the pmoA and pmoB subunits. The subunits interact with each other via their alpha helices as well as through a short helix and beta hairpin from pmoA which interact with the soluble region of pmoB (Lieberman and Rosenzweig, 2005) (Figure 3.13 A). Based on data from PDBePISA (www.ebi.ac.uk/pdbe/pisa), pmoB and pmoA interact via 13 hydrogen bonds and 3 salt bridges on their interface. PmoB and pmoC are instead more loosely associated, with a sole point of contact between short periplasmic pmoC loops and the pmoB copper site (Figure 3.13 B). Data from PDBePISA shows 2 hydrogen bonds and 1 salt bridge between the pmoB and pmoC interfaces.

To test whether the native interactions would cause pmoB to relocate to the ER, pmoB-GFP was co-expressed with pmoA-RFP in tobacco cells. The location of fluorescent tags should not compromise any potential interactions as fluorophores were placed in sterically available termini, as indicated with orange arrows in Figure 3.13. Additionally, an 11 amino acid linker sequence consisting of the Gateway recombination site is present between the proteins and fluorophores, which would allow flexibility during folding.

However, transient co-expression of pmoB-GFP and pmoA-RFP showed that pmoB did not relocate to the ER and instead maintained its phenotype forming ring structures with a cytosolic background (Figure 3.14). This highlighted the need to engineer the protein sequence in order to target pmoB to the ER.

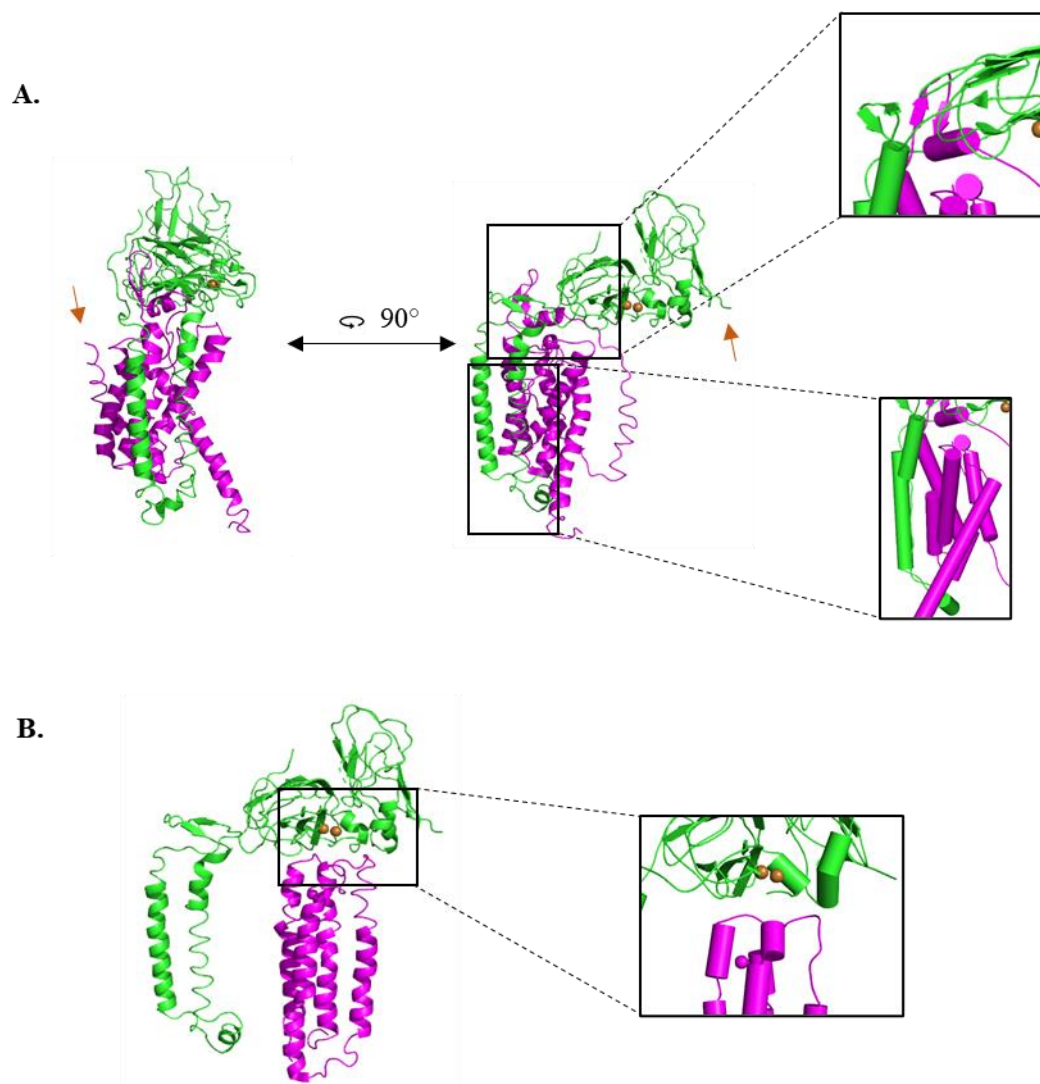


Figure 3.13. Points of contact between pmoB and other subunits.

A) Structures of pmoB (green) and pmoA (magenta) within the pMMO complex. The two interfaces are enlarged showing an association between a periplasmic region of pmoB and a short helix and beta hairpin of pmoA (top), and between the alpha helices (bottom). Orange arrows point towards the C-terminus of the subunits where fluorophores were fused. B) Structures of pmoB (green) and pmoC (magenta), showing the enlarged interface.

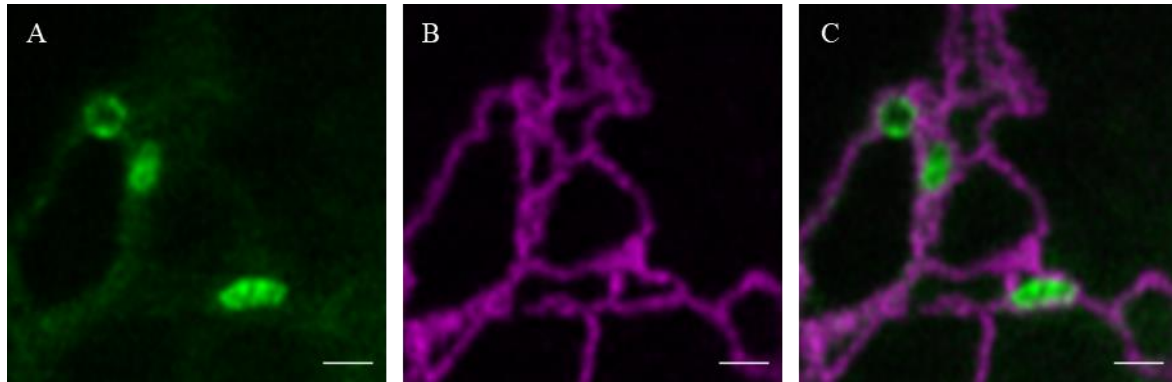


Figure 3.14. Co-expression of pmoB-GFP and pmoA-RFP.

Confocal image showing the localization of pmoB-GFP (A), pmoA-RFP (B), and the merged channel (C). Scale bar = 1 μm .

3.2.3 Targeting pmoB to the ER

3.2.3.1 Targeting pmoB via a plant signal peptide is successful but results in some protein escape to the vacuole

Co-expression of pmoA and pmoB was not sufficient to achieve an ER localization of pmoB. It was therefore necessary to re-engineer pmoB with the aim of expressing all three recombinant subunits in the same membrane compartment. Thus, the bacterial signal peptide of pmoB was replaced with the plant signal peptide from pumpkin 2S albumin. The signal peptide from pumpkin 2S albumin is reported to direct the seed storage protein 2S albumin to the ER before its cleavage in the ER lumen between residues A-Y (Hara-Hishimura *et al.*, 1993) (Figure 3.15). This signal peptide has been used previously to produce a chimeric GFP protein (spGFP-HDEL) which labels the ER network (Mitsuhashi *et al.*, 2000), referred to in this study as GFP-HDEL, and has been employed in several studies as an ER marker.



Figure 3.15. Signal peptide from pumpkin 2S albumin.

Sequence of the signal peptide from pumpkin 2S albumin employed in this study for targeting of the pmoB subunit, showing the cleavage site with a blue dotted line. Sequence retrieved from (Hara-Hishimura *et al.*, 1993).

A construct was cloned using overlap PCR to add the plant signal peptide N-terminally to Btrunc1, yielding “spBtrunc-GFP”. The engineered protein spBtrunc-GFP was successfully targeted to the ER, as seen by transient expression in tobacco cells (Figure 3.16 A). However, expression levels of this protein were very low, this can be evidenced by the presence of background autofluorescence in the cell borders, and autofluorescence in stomata, which have a signal intensity comparable to that present in the ER network (Figure 3.16 A and B). Expression levels of GFP-pmoA are shown as a comparison in lower magnification level ROIs including multiple cells (Figure 3.16 C). For clarity, all the vectors constructed for targeting are summarized in appendix 3 (section 9.2), indicating their subcellular localization and expression levels.

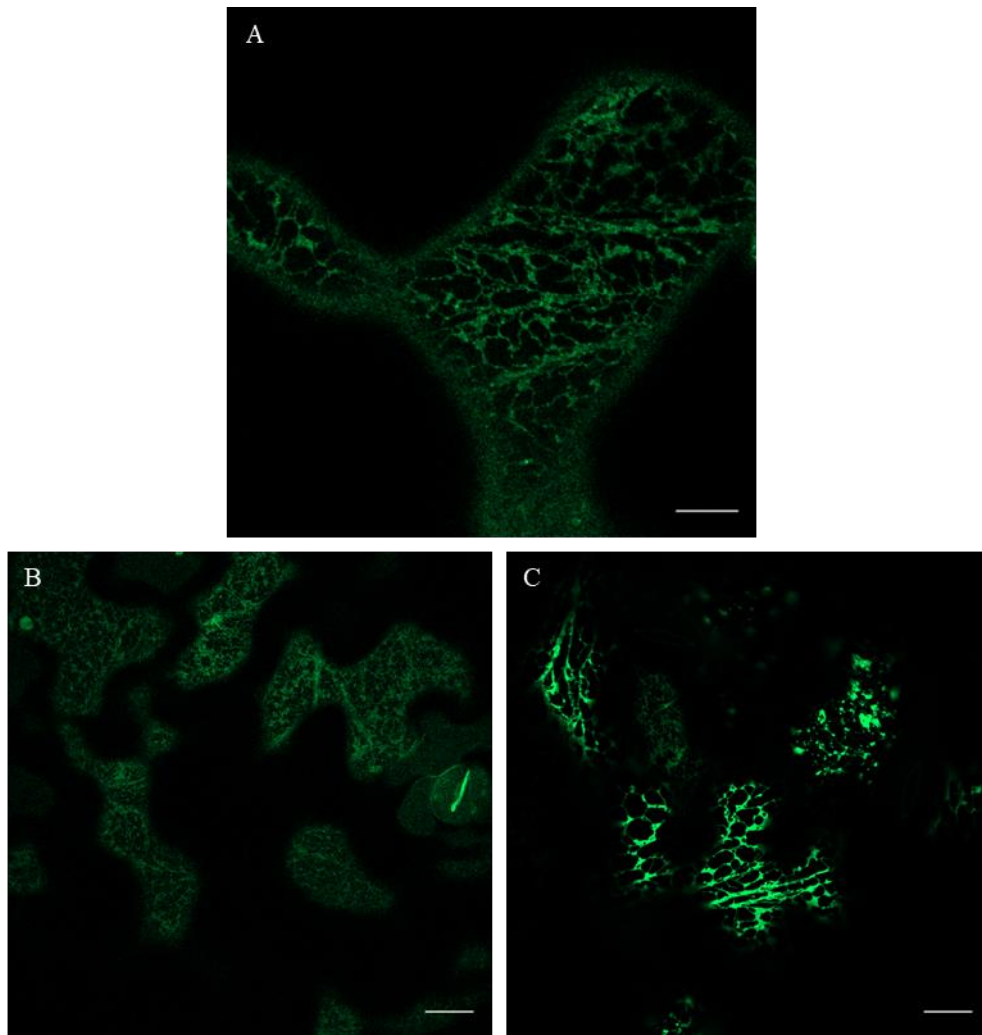


Figure 3.16. Expression of spBtrunc-GFP.

SpBtrunc-GFP is successfully targeted to the ER (A). Lower magnification ROIs show that expression levels of spBtrunc-GFP (B) are low compared to GFP-pmoA (C). Scale bar = 10 μm (A) and 20 μm (B and C).

One explanation for the low fluorescence of spBtrunc-GFP may be due to the protein escaping the ER and being trafficked to downstream compartments in the secretory pathway. Specifically, localization of GFP in the apoplast or vacuole would result in low or null levels of fluorescence due to the acidic pH and proteolysis (Batoko *et al.*, 2000; Tamura *et al.*, 2003; Samalova *et al.*, 2006).

To test the presence of spBtrunc in the apoplast or vacuole a fusion protein was created with an RFP moiety (“spBtrunc-RFP”), which can withstand the conditions in the vacuole and

apoplast and can be detected in these compartments. SpBtrunc-RFP localized both to the vacuole and to the ER at 2 days after infiltration. Figure 3.17 A-C shows the dual localization through 3D Z-stacks of the same group of cells. At a lower focal plane the vacuolar signal was seen and ER localization could only be observed as an irregular bright signal in the cell borders (Figure 3.17 A). At a mid-focal plane some cells still displayed vacuolar signal whilst others had reached the top of the cell and showed the ER as a reticulate network (Figure 3.17 B). At a higher focal plane most cells of the ROI were now focused at the top, where only the ER network is visible (Figure 3.17 C). At 3 days after infiltration, localization was entirely vacuolar (Figure 3.17 D), indicating that any spBtrunc retention in the ER was lost and the protein was trafficked fully to the vacuole.

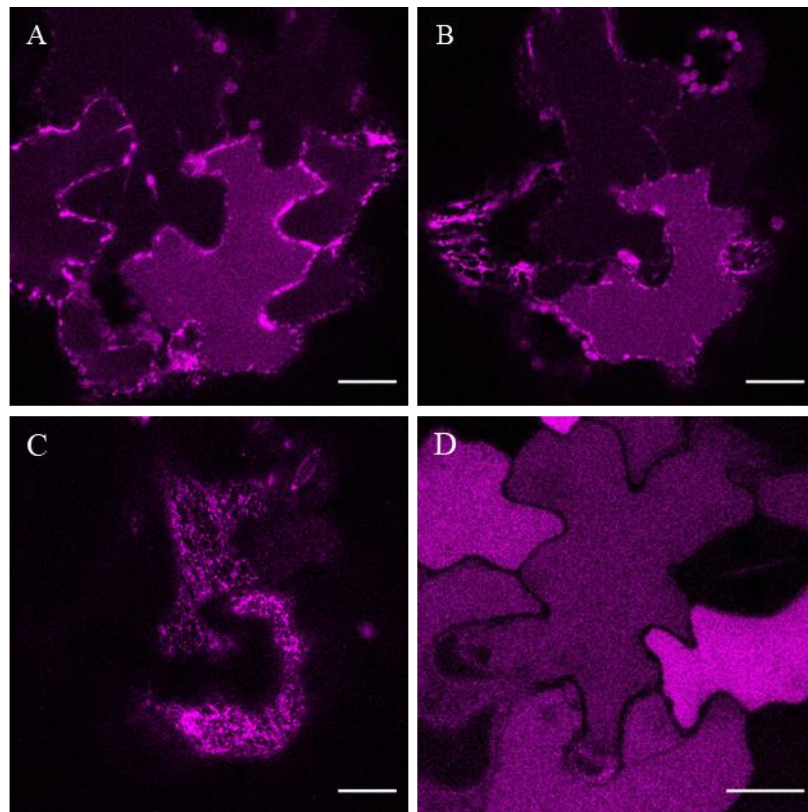


Figure 3.17. Subcellular localization of spBtrunc-RFP.

Transient expression of spBtrunc-RFP in tobacco cells shows dual localization in the ER and the vacuole at 2 days after infiltration (A-B: Z-stacks of the same ROI), and entirely vacuolar localization at 3 days after infiltration (D). Scale bar = 20 μ m.

3.2.3.2 Addition of the HDEL retention signal produces accumulation of spBtrunc-RFP in the ER

Given the partial vacuolar localization of spBtrunc-RFP at 2 days after infiltration, and complete vacuolar localization at 3 days after infiltration, the addition of ER retention motifs was investigated to enhance ER accumulation of the protein. Two constructs were cloned using C-terminal retention motifs as shown in Figure 3.18 A. The HDEL tag was predicted to aid retention assuming that the C-terminus of spBtrunc was localized to the ER lumen, as expected from the topology on the crystal structure (see Figure 3.9). A construct with the KKRY tag was also constructed to test whether the recombinant protein was misfolded and the C-terminus was facing the cytosol.

Transient expression assays showed that spBtrunc-RFP-HDEL achieved strong ER retention at 2 days after infiltration (Figure 3.18 B and C). Moreover, the expression levels were high and no visible alterations to ER morphology were observed. At 3 days after infiltration, expression levels were lower, and a very faint vacuolar localization could be detected in some cells (Figure 3.18 D) in addition to cells which had only ER localization (Figure 3.18 E). This showed a marked difference compared to spBtrunc-RFP at 3 days which was completely vacuolar (Figure 3.18 D).

In contrast, spBtrunc-RFP-KKRY behaved in a similar manner to spBtrunc-RFP, i.e. the protein was present both in the vacuole and the ER at 2 days after infiltration (Figure 3.18 F) and was fully vacuolar at 3 days (Figure 3.18 G), indicating that the protein may be correctly folded with a luminal C-terminus. The ER morphology at 2 days was also severely altered showing a diffuse network and strong aggregate-like lumps (Figure 3.18 F).

These findings revealed that replacing the bacterial signal peptide for a plant signal peptide and adding a C-terminal HDEL motif was a successful strategy to accumulate pmoB in the ER membrane. Moreover, the C-terminus of pmoB appears to be in the ER lumen, since retention is achieved via HDEL and not the dilysine motif, providing some evidence towards correct folding of the recombinant protein in eukaryotic cells.

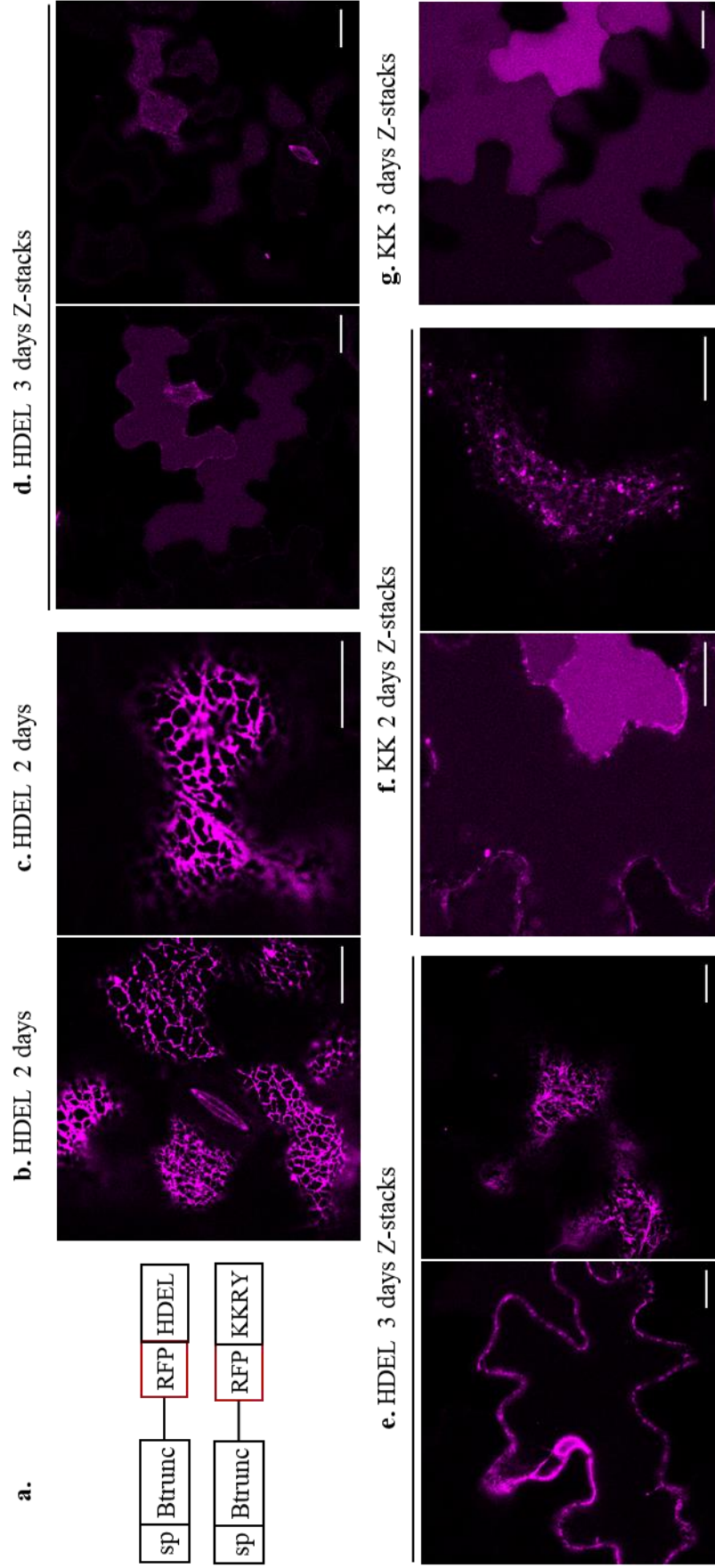


Figure 3.18. Retention of ER localised pmoB (spBtrunc-RFP) via the addition of C-terminal motifs.

Schematic representation of the constructs used (a), expression of spBtrunc-RFP-HDEL at 2 days after infiltration showing high expression at lower and higher magnification (b) and (c), expression of spBtrunc-RFP-HDEL at 3 days after infiltration showing low vacuolar signal (d) and ER localization (e), expression of spBtrunc-RFP-KKRY at 2 days after infiltration showing dual vacuolar and ER localization (f) and at 3 days showing only vacuolar localization (g). Scale bar = 20 μm .

3.2.3.3 The HDEL retention signal may marginally increase fluorescence of spBtrunc-GFP

To confirm vacuolar localization as the reason for lower GFP intensity levels of the spBtrunc-GFP construct, the HDEL motif was now added to the GFP fusion protein, obtaining “spBtrunc-GFP-HDEL”. Surprisingly, intensity levels did not increase in great measure with the addition of the retention sequence. Some cells appeared to display slightly stronger expression levels, however, small differences in signal intensity are difficult to evaluate given the unevenness in expression levels within different regions of interest (ROIs) and varying expression levels between plants.

Thus it appeared that spBtrunc-GFP-HDEL was not retained as well as spBtrunc-RFP-HDEL. This could potentially be due to a difference in folding of the two fusion proteins leading to reduced ability of HDEL to bind to ERD2 in the GFP version. However, the fluorophores used (EGFP and mRFP1) possess very similar 3D structures. An alignment of the protein structures produced in PyMOL is shown in Figure 3.19 A. The structure of mCherry was used given that the structure of mRFP1 is not available, but both RFP proteins share a high sequence similarity (Figure 3.19 B). The C-terminus of the GFP and RFP proteins would appear to be sterically available in both cases. One difference that may be relevant is that EGFP has the ability to form weak dimers, whereas mRFP1 possesses mutations to make it strictly monomeric. Dimerization of EGFP could reduce accessibility of the HDEL tag.

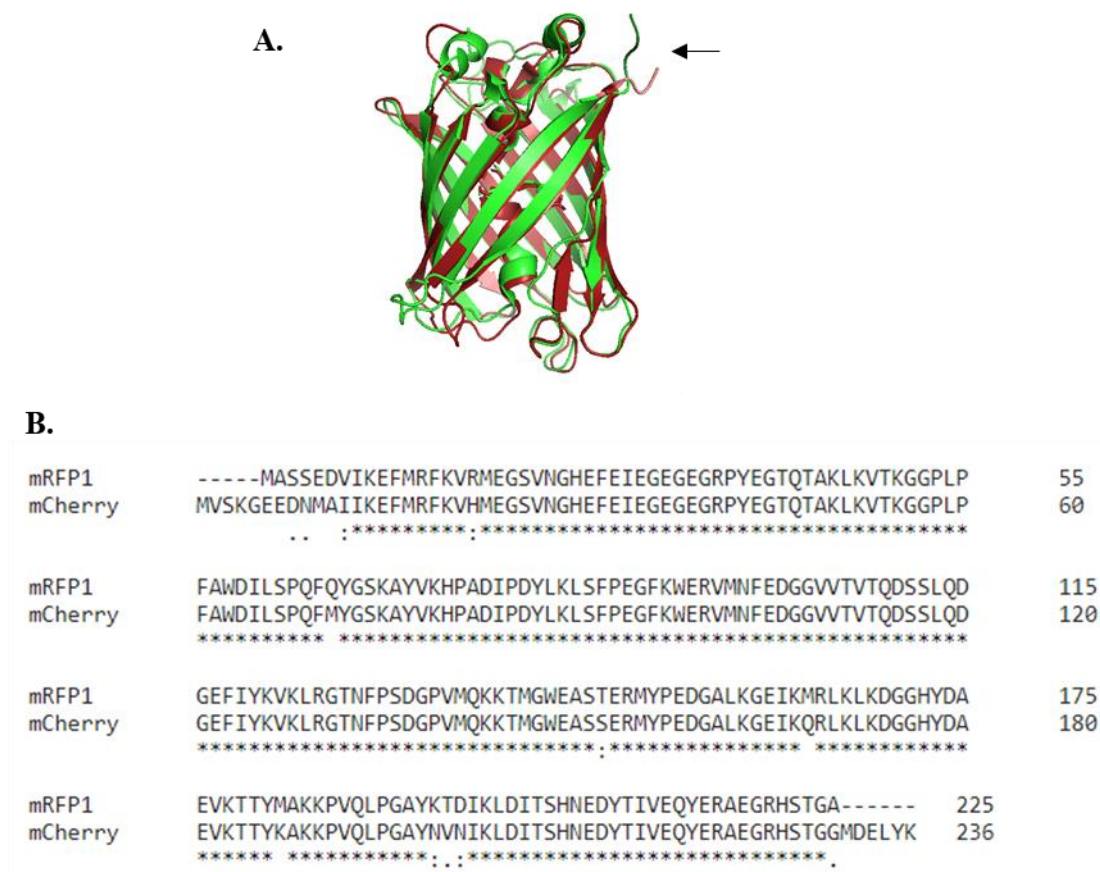


Figure 3.19. Structure of fluorescent proteins.

A) Alignment of protein structures of EGFP (green) and mCherry (red) using PyMOL, the C-terminus is signaled with a black arrow. B) Alignment of protein sequences of mCherry and mRFP1, showing high sequence similarity using Clustal Omega.

A method was established in order to quantitatively describe and compare expression levels of different tagged versions using GFP-HDEL and autofluorescence levels as references (see materials and methods). The results from the quantification of fluorescence intensity show the mean and standard error of three biological replicates of all GFP tagged constructs (Figure 3.20). Expression levels of GFP-pmoA and GFP-pmoC were comparable to expression levels of GFP-HDEL. However, much lower expression levels were seen for the spBtrunc-GFP construct, which appears only slightly higher than autofluorescence levels. Addition of the HDEL tag to the ER targeted pmoB subunit (spBtrunc-GFP_{HDEL}) produced marginally higher expression levels. However, those did not reach the amounts obtained by the other pMMO subunits. To try to make the HDEL tag more available and reduce any effects of potential

dimerization of EGFP, the addition of a modified tag “HDEL2” was investigated (sequence: GGGHHHHHHDEL). HDEL2 consists of the exact amino acid sequence present in the GFP-HDEL marker (Mitsunishi *et al.*, 2000), and contains a flexible glycine linker, as well as repeated histidine residues. The HDEL2 tag would hypothetically allow more flexibility and better binding to the ERD2 receptor. However, transient expression of spBtrunc-GFP-HDEL2 yielded similar results to spBtrunc-GFP-HDEL, with no marked increase in fluorescence.

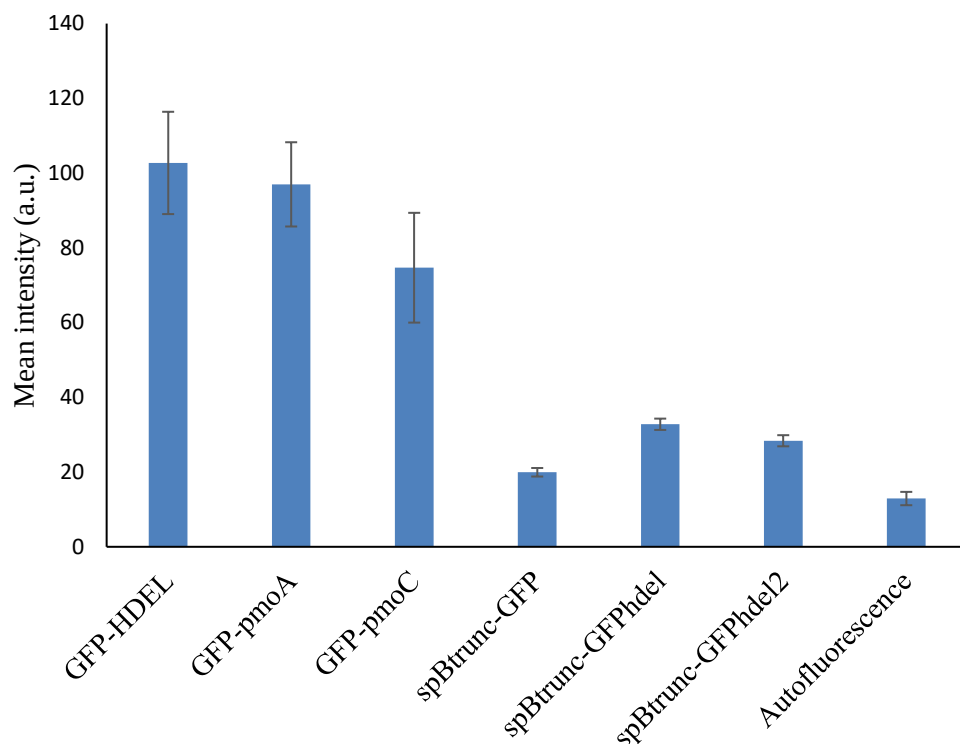


Figure 3.20. Fluorescence intensity of GFP-tagged subunits.

The average fluorescence intensity of GFP was measured in 20 ROIs (technical repeats) and 3 plants (biological replicates) for each construct, under constant imaging conditions. All constructs within a biological replicate were infiltrated on the same leaf. Intensity levels were obtained from areas within the ROI that fell above an automated threshold level. Autofluorescence was measured on areas which were not infiltrated. The means of three biological replicates with the standard errors are shown.

3.2.3.4 Suppression of post-transcriptional gene silencing increases expression levels

To further investigate the reasons for the low fluorescence levels of spBtrunc-GFP and spBtrunc-GFP_{hdel} the subunits were co-expressed with the P19 protein, a suppressor of post-transcriptional gene silencing (Voinnet *et al.*, 2003), which has often been used in *Nicotiana benthamiana* to improve expression levels of recombinant proteins. The expression levels of spBtrunc-GFP and spBtrunc-GFP_{hdel} in *Nicotiana benthamiana* at 2 days after infiltration were very low, as expected, however, expression levels of both constructs improved upon co-expression with p19 (Figure 3.21). The same plant was imaged again at 3 days after infiltration to assess the evolution of protein accumulation. After 3 days, expression of the constructs in the absence of p19 was lower than at 2 days, and the addition of p19 induced an increase in protein levels of spBtrunc-GFP to similar levels of those at 2 days. The addition of p19 to spBtrunc-GFP_{hdel} increased protein accumulation in the ER to a higher extent at 3 days.

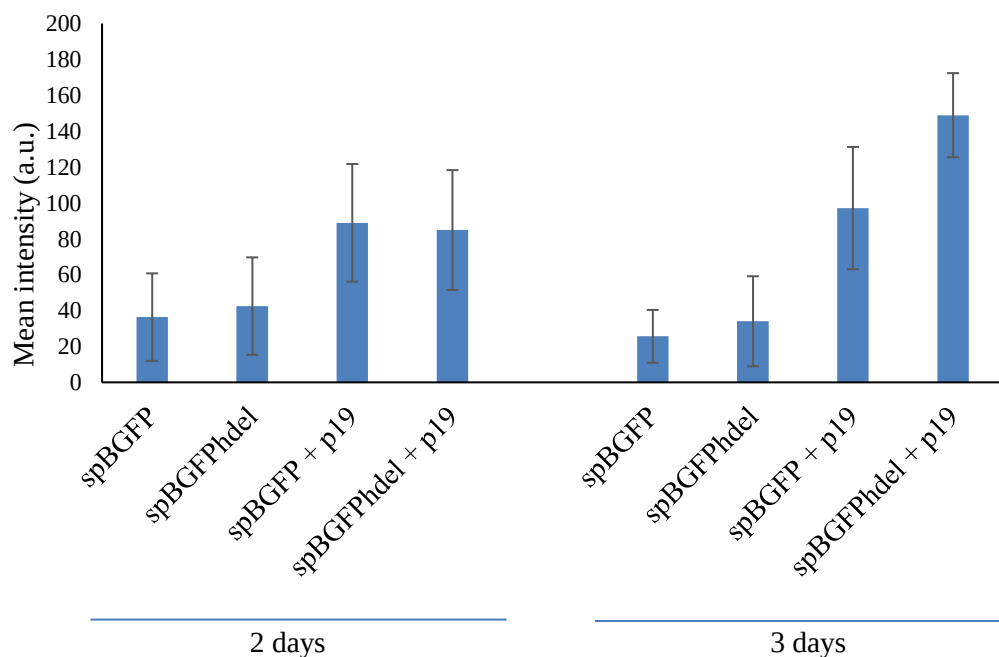


Figure 3.21. Co-expression of spBtrunc-GFP and spBtrunc-GFP_{hdel} with P19.

Fluorescence intensity levels are shown for cells expressing spBtrunc-GFP and spBtrunc-GFP_{hdel} in absence and in the presence of P19. The mean intensity of 20 ROIs is shown with the standard deviation (1 biological replicate). Expression was measured at 2 days and 3 days after infiltration in *Nicotiana benthamiana*.

3.2.4 Testing protein-protein interactions

3.2.4.1 Co-expression of pMMO subunits on the ER

Previously, all of the bacterial pMMO subunits were localized individually to the ER. In order to determine if more than one subunit could be localized to the ER simultaneously, constructs with fluorescent protein tags were co-infiltrated into tobacco plants. PmoC and pmoA were both infiltrated at an OD₆₀₀ of 0.05 due to the high levels of expression of these constructs and to mitigate the damage observed with expression of pmoC, whereas the ER localized pmoB (spBtrunc-RFP_{hdel}) was infiltrated at OD₆₀₀ 0.1. The randomness of T-DNA insertions following infiltration with agrobacterium resulted in different protein stoichiometries between cells (Figure 3.22, merged channel). Co-expression of pmoA and pmoC showed that both subunits jointly localized to the ER, but with most cells showing disruption of the ER network, whereas cells with only GFP-pmoA showed an intact ER network (Figure 3.22 A). Co-expression of pmoA with pmoB resulted in both subunits localizing to the ER without visible damage to the ER network (Figure 3.22 B). Finally, the co-expression of pmoC and pmoB showed that both subunits localized to the ER with elevated levels of ER disruption, except in cells with faint GFP-pmoC expression where less ER disruption was observed (Figure 3.22 C). It can be inferred that the ER disruption is caused by the presence of the pmoC subunit, in accordance with the results from single infiltrations.

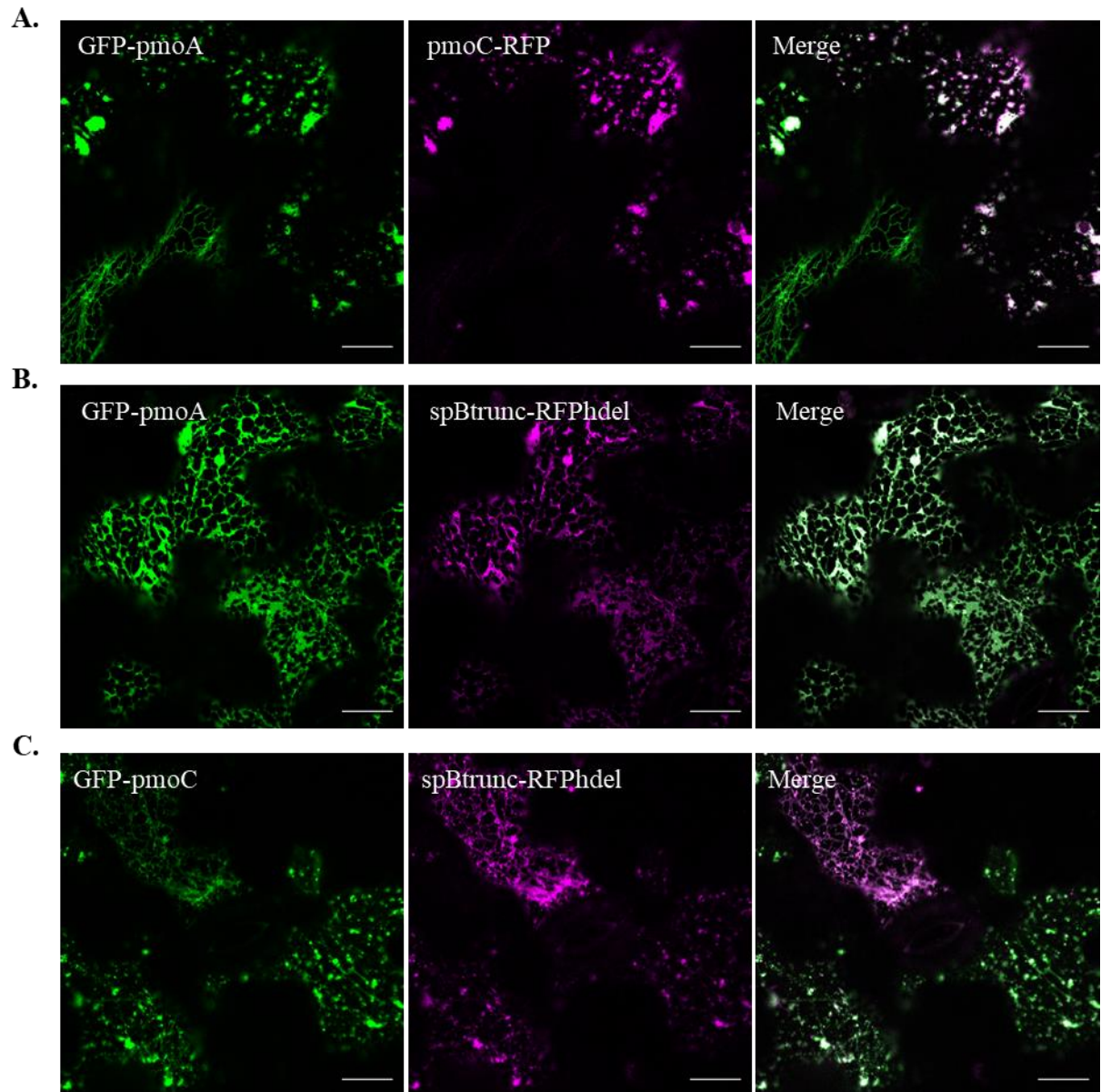


Figure 3.22. Pairwise co-expression of pMMO subunits.

Co-expression of the subunits tagged with green or red fluorescent proteins shows simultaneous localization in the ER when expressed in a pairwise manner, and ER disruption in the presence of pmoC. Scale bar = 20 μ m.

3.2.4.2 Determination of protein-protein interactions with FRET-FLIM

PMMO is a nonamer with three repetitions of the pmoC, pmoA and pmoB trimer. Successful reconstitution of the complex requires native interactions to form in the plant, mainly between pmoC and pmoA, and between pmoA and pmoB. Having previously managed to co-express all three pMMO subunits in the ER in a pairwise manner, FRET-FLIM was used to assess the

presence of protein-protein interactions on the ER membrane. The occurrence of FRET is dependent on a separation of less than 10 nm between the fluorophores, with FRET efficiency decreasing as the distance between the molecules increases up to 10 nm. The chromophores are found buried in the centre of the fluorophore beta barrel structures, which have a diameter of around 2.4 nm, reducing the effective distance for FRET to less than 7 nm (Bajar *et al.*, 2016). Given the thickness of the ER membrane being approximately 4 nm (Mitra *et al.*, 2004), it was considered unlikely that fluorophores on different side of the membrane would present a significant interaction and therefore the expected topology of the pMMO proteins was considered in the experimental design.

GFP-pmoA and pmoA-GFP were used as donors to test interactions with both termini of pmoC (mCherry-pmoC and pmoC-RFP) and with the C-terminus of pmoB (spBtrunc-RFPhdel) (Figure 3.23). The N-terminus of pmoB was not fused to a fluorophore and therefore not included in this study because of the requirement of the N-terminal signal peptide. In order to test the remaining pairwise interaction between pmoC and pmoB, GFP-pmoC was used as the donor and spBtrunc-RFPhdel as the acceptor (Figure 3.23). Additionally, interactions of the subunits with themselves (i.e. by dimerization) were tested.

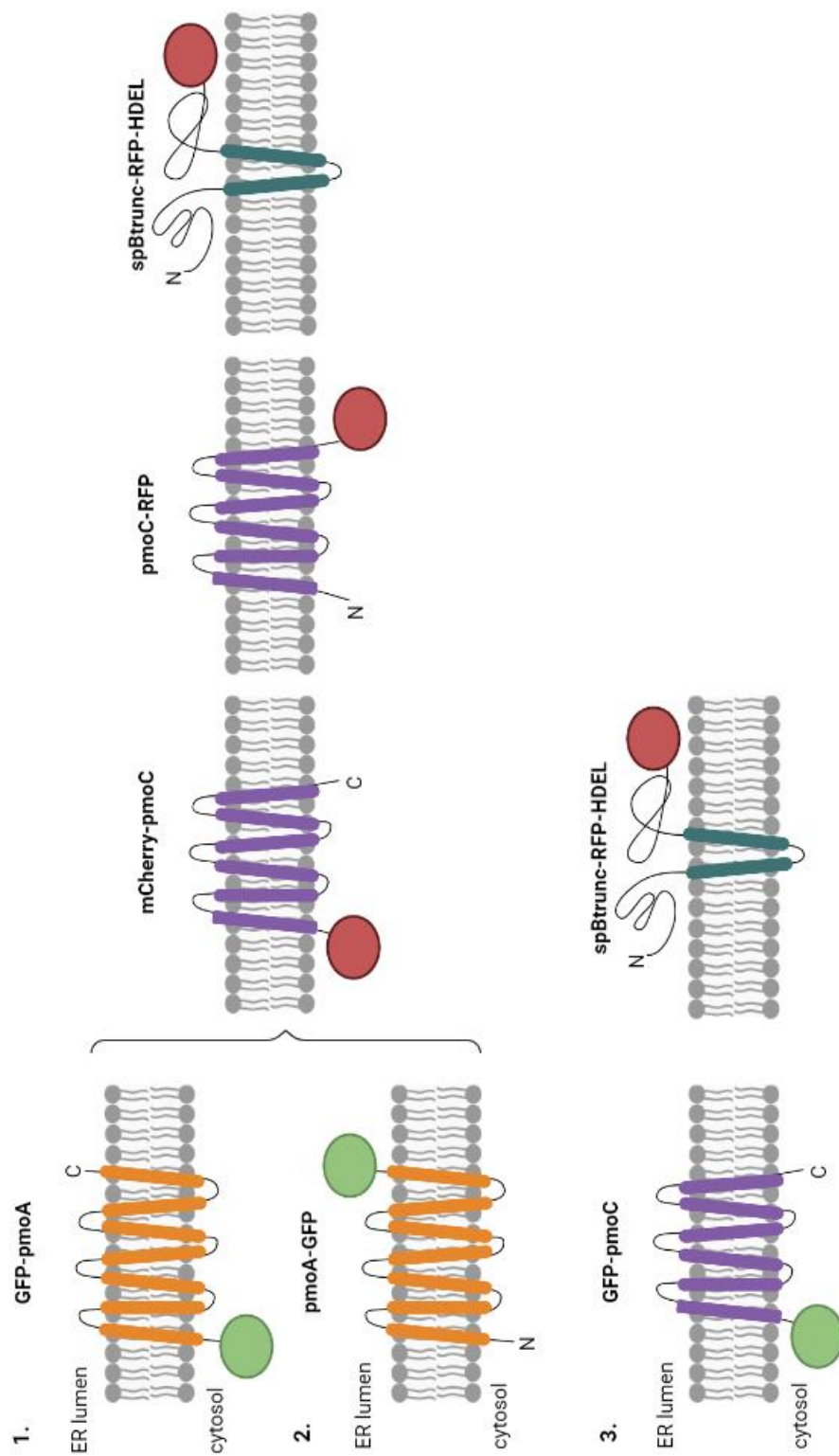


Figure 3.23. FRET-FLIM design.

Design of FRET-FLIM interaction partners and their expected topology. Donor molecules are fused to GFP and acceptor molecules are fused to RFP. Fluorophores localized on opposite site of the membrane are not expected to interact via FRET.

As the ER is a highly dynamic organelle, FRET-FLIM was carried out in the nuclear envelope, which is contiguous with the ER membrane but has a relatively low mobility, to give more reliable measurements (Sparkes *et al.*, 2010; Kriechbaumer *et al.*, 2015). The lifetime of the GFP tagged subunit (donor) was measured in the absence and in the presence of an RFP tagged subunit (acceptor). A reduction in the average lifetime of the donor in the presence of an acceptor was indicative of a protein-protein interaction (see section 3.1.3). An example image of GFP-pmoA without an interaction partner and with an interaction partner (mCherry-pmoC) is shown in Figure 3.24. A region of the nuclear envelope was selected (Figure 3.24 A), and lifetime decay curves were plotted for every pixel and fitted to an exponential curve (Figure 3.24 C). Curves which showed a good fit (where the chi-squared fell within the range of 0.9 and 1.2) were selected. Color-coded maps show the distribution of the lifetime values obtained for the region of interest in picoseconds, with blue colors indicating a longer lifetime (Figure 3.24 B). Example images of the remaining combinations tested are shown in Supplementary Figure 3.2.

The average lifetime values of the GFP-tagged donor molecules alone ranged between 2.34 and 2.45 ns (Table 3.3). A significant reduction in the lifetime of GFP-pmoA was found in the presence of mCherry-pmoC and pmoC-RFP, indicating a likely interaction between the N-terminus of pmoA and both termini of pmoC (Figure 3.25). The same pattern of interactions was observed for pmoA-GFP, showing that both termini of pmoA and pmoC interact, whereas no interaction between pmoA and spBtrunc was found. In addition, the reduction in lifetime of GFP-pmoA with pmoA-RFP suggests that the subunit interacts with itself in the plant. Confirming the affinity between pmoA and pmoC, a significant reduction in the GFP-pmoC lifetime was found in the presence of pmoA-RFP. The GFP-pmoC lifetime in the presence of spBtrunc-RFP_{hdel} appears in the interaction region of the chart due to the low standard error in these samples (Figure 3.25), however, with a lifetime reduction of just 0.05 ns this was not considered indicative of an interaction. The pmoC subunit also appears to interact with itself.

Finally, a significant reduction in the spBtrunc-GFP_{hdel} in the presence of spBtrunc-RFP_{hdel} indicated a likely interaction of spBtrunc with itself.

This interaction pattern was contrary in part to what was expected in the case of the predicted topology of pmoA (Figure 3.26 A), where the C-terminus of pmoA would have been expected to interact with the C-terminus of spBtrunc due to both being localized to the ER lumen. One explanation for the pattern of interactions could be that the C-terminus of pmoA localizes to the cytosol instead of to the ER lumen, explaining the interactions with the pmoC termini and lack of interaction with pmoB (Figure 3.26 B).

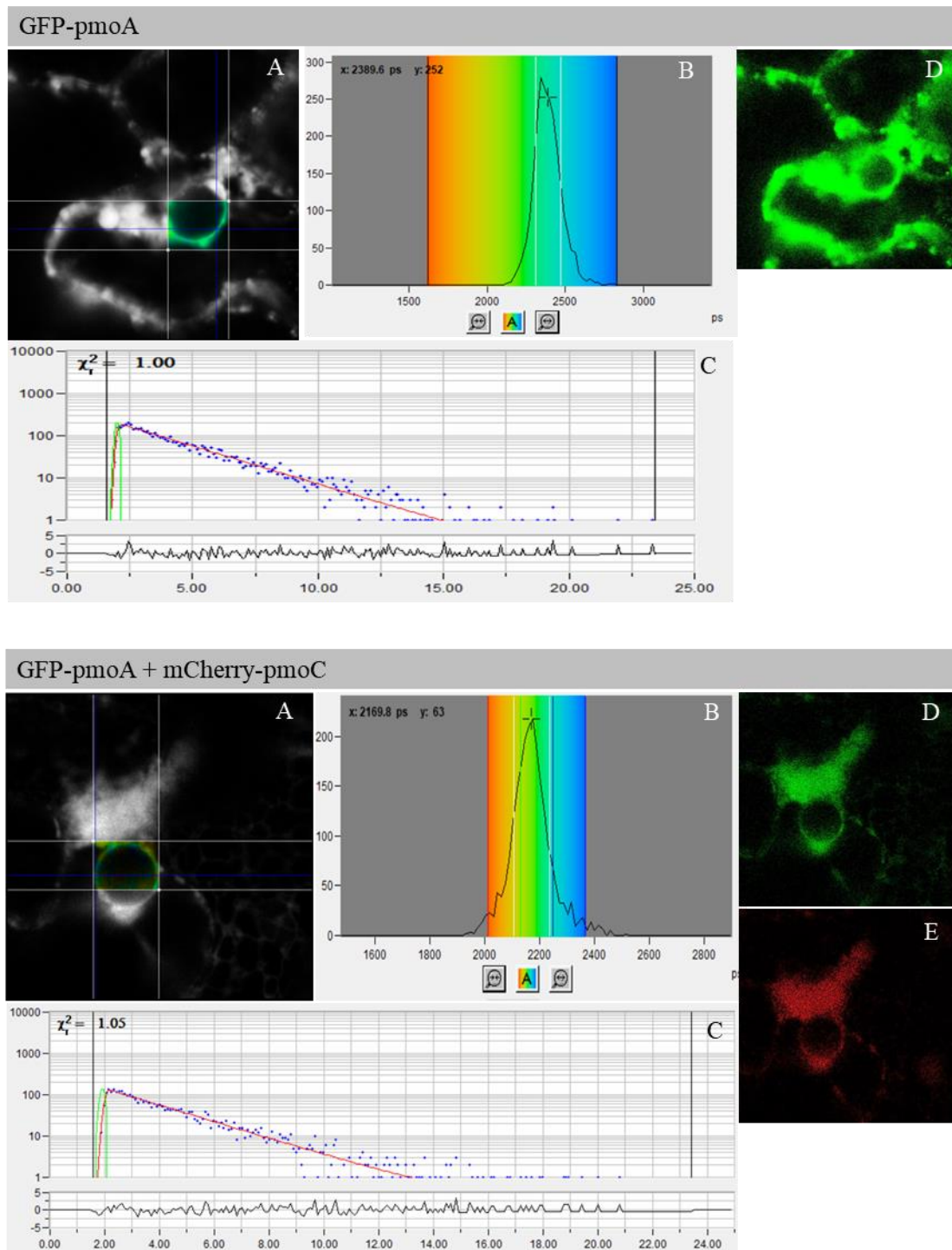


Figure 3.24. Example images and processing used for FRET-FLIM for GFP-pmoA alone and with an interaction partner.

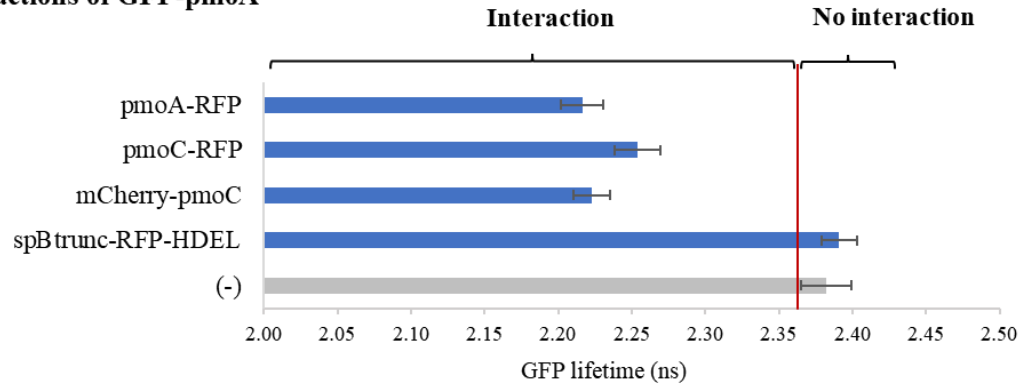
(A) Pseudo-colored lifetime map showing the region of interest of the nuclear envelope analyzed, (B) color-coded distribution of the GFP lifetime in the region of interest, (C) lifetime decay curve with an exponential fit of a pixel within the studied region, (D, E) raw FRET-FLIM images.

Table 3.3. GFP lifetime values obtained from FRET-FLIM.

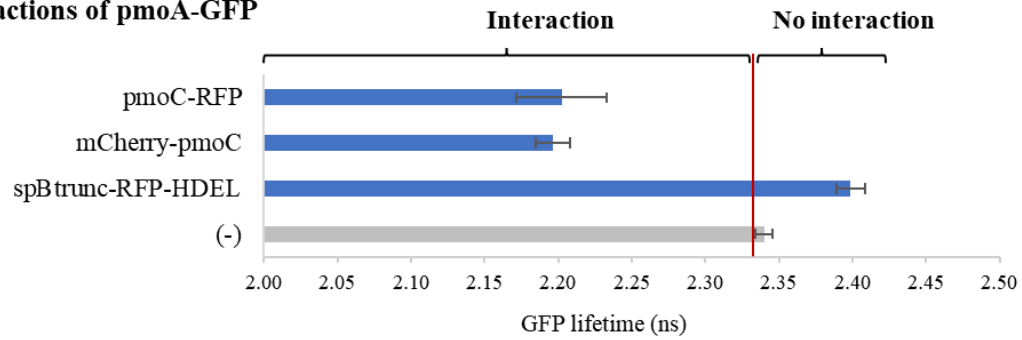
The average lifetime (nanoseconds) is shown for the different donor molecules with and without an acceptor molecule. Average values, the standard error (SE), and the number of cells (n) are shown, as well as the difference in lifetime between the donor alone and in the presence of and RFP-tagged construct (Δ). The cells were taken from 3-4 different plants in each case.

Donor	Acceptor	Average (ns)	SE	n	Δ (ns)
GFP-pmoA	-	2.38	0.02	10	
GFP-pmoA	spBtrunc-RFPhdel	2.39	0.01	10	+0.01
GFP-pmoA	mCherry-pmoC	2.22	0.01	8	-0.16
GFP-pmoA	pmoC-RFP	2.25	0.02	8	-0.13
GFP-pmoA	pmoA-RFP	2.22	0.01	8	-0.17
pmoA-GFP	-	2.34	0.01	8	
pmoA-GFP	spBtrunc-RFPhdel	2.40	0.01	8	+0.06
pmoA-GFP	mCherry-pmoC	2.20	0.01	8	-0.14
pmoA-GFP	pmoC-RFP	2.20	0.03	8	-0.14
GFP-pmoC	-	2.42	0.01	10	
GFP-pmoC	spBtrunc-RFPhdel	2.37	0.01	12	-0.05
GFP-pmoC	pmoA-RFP	2.16	0.03	8	-0.26
GFP-pmoC	mCherry-pmoC	2.19	0.02	8	-0.23
spBtrunc-GFPhdel	-	2.45	0.01	6	
spBtrunc-GFPhdel	spBtrunc-RFPhdel	2.31	0.02	6	-0.15

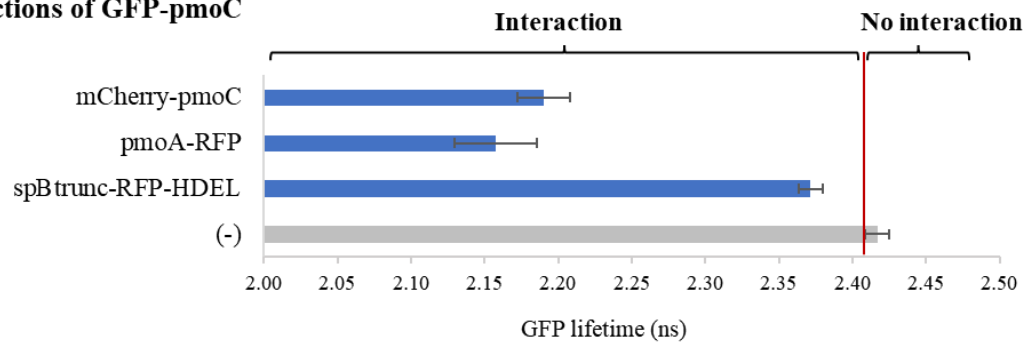
Interactions of GFP-pmoA



Interactions of pmoA-GFP



Interactions of GFP-pmoC



Interactions of spBtrunc-GFP_{HDEL}

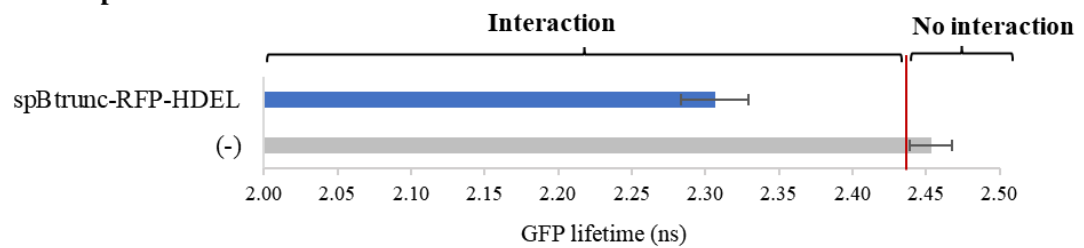


Figure 3.25. GFP lifetimes and interactions between donor and acceptor molecules in FRET-FLIM.

Lifetime values with their standard deviation are shown. Significant reductions in the GFP lifetime of the donor (-) are classified as interactions.

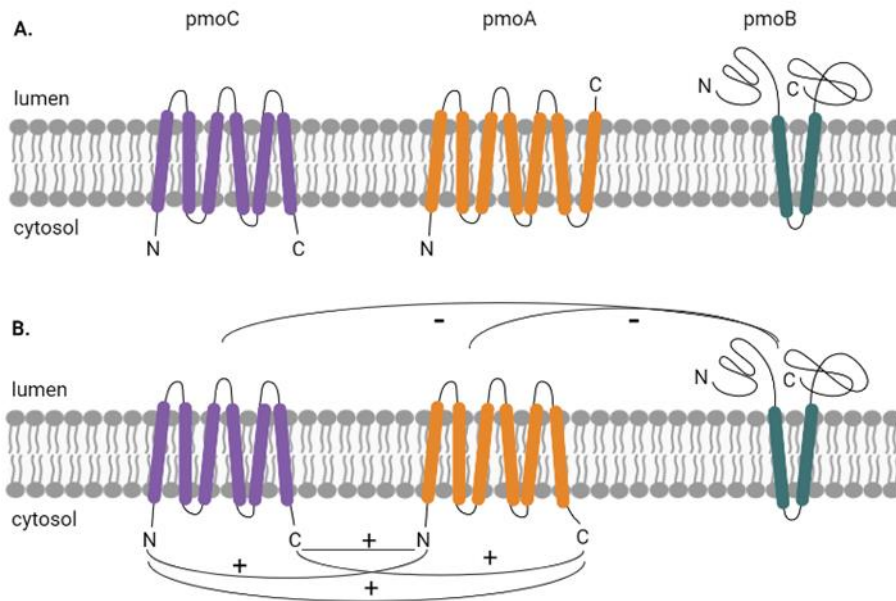


Figure 3.26. Diagram of topology and FRET-FLIM interactions on the plant ER.

(A) Expected topology of pMMO proteins on the ER membrane, (B) proposed topology on the ER membrane based on FRET-FLIM interactions (+ = interaction, - = no interaction).

3.2.4.3 ER retention of vacuolar pmoB in the presence of pmoA

FRET-FLIM studies showed an interaction between the pmoC and pmoA subunits, in accordance with what was expected due to the close association of both subunits in the crystal structure. No interaction was detected between pmoA and pmoB or between pmoC and pmoB, even though the pmoA and pmoB subunits interact tightly in the crystal structure. It is possible that an interaction between these subunits via FRET-FLIM could not be detected due to the fluorophores being on different sides of the membrane (i.e. pmoA termini residing in the cytosol and pmoB in the ER lumen). Therefore, an alternative microscopy-based method was used to indirectly test this interaction. The spBtrunc-RFP construct localized to the ER and vacuole at 2 days after infiltration, and solely to the vacuole at 3 days after infiltration (see Figure 3.17). When spBtrunc-RFP was co-expressed with clover-pmoA, spBtrunc-RFP was now partially retained in the ER at three days after infiltration (Figure 3.27). As a control, an ER membrane marker (GFP-Calnexin) was also co-expressed with spBtrunc-RFP and no retention was observed in this case. From 106 cells counted in 30 ROIs which co-expressed

spBtrunc-RFP and clover-pmoA at detectable levels, 49% of cells showed retention in the ER, whereas 100% of cells co-expressed with GFP-Calnexin showed localization of spBtrunc-RFP only in the vacuole. These results suggested that pmoA might be interacting with pmoB and causing the observed ER retention of spBtrunc.

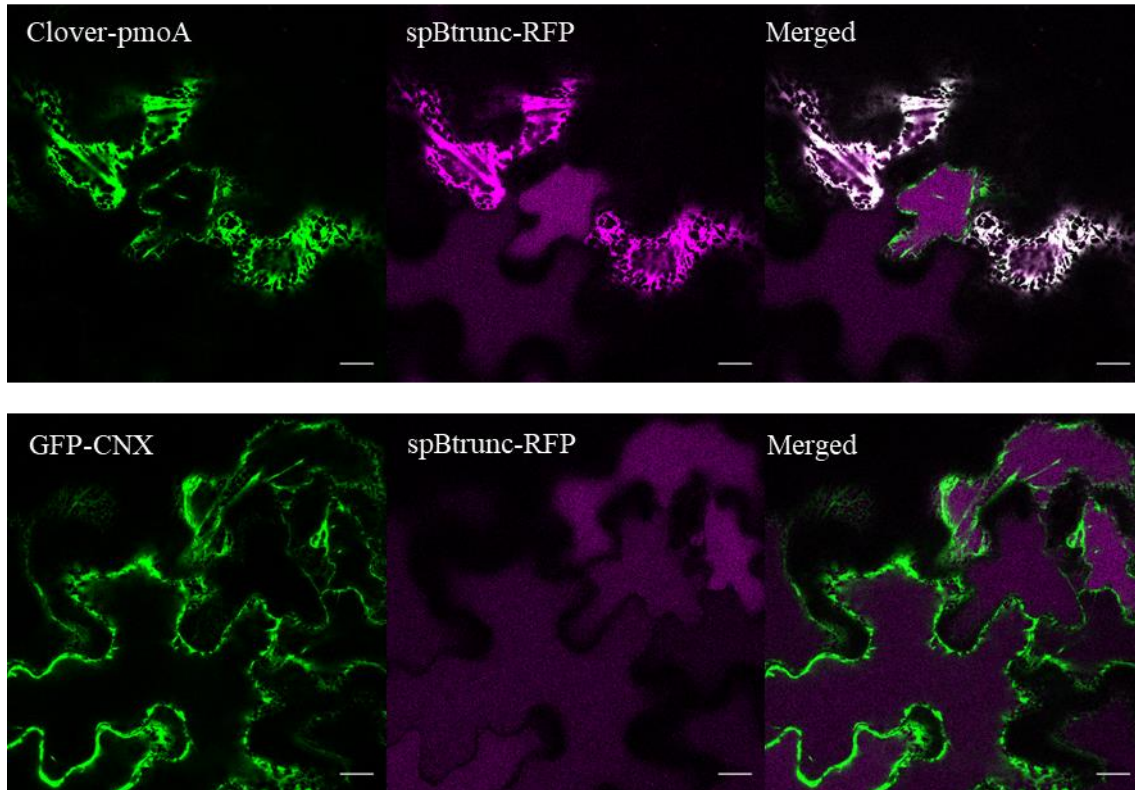


Figure 3.27. Co-expression between spBtrunc-RFP and Clover-pmoA.

Co-expression of spBtrunc-RFP and Clover-pmoA induced ER retention in approximately half of the cells at three days after infiltration. Co-expression of spBtrunc-RFP with GFP-Calnexin (GFP-CNX) did not cause ER retention, instead showing 100% vacuolar localization in the red channel. Scale bar = 10 μ m.

3.2.5 Topology of pMMO proteins on the ER membrane

The results from protein-protein interaction studies with FRET-FLIM as well as the ER retention of pmoB with HDEL suggest a slight difference between the topology expected from the crystal structure and the actual topology of pMMO proteins in the ER membrane. Results so far point towards pmoA having both termini facing the cytosol, as opposed to the crystal

structure showing the N-terminus of pmoA facing the cytosol and the C-terminus of pmoA localized to the periplasmic space in bacteria.

In order to test the topology of the subunits, a redox-based method was used based on fusion to the redox-sensitive GFP protein (roGFP2) on either protein terminus (Hanson *et al.*, 2004; Brach *et al.*, 2009). The following constructs were cloned and transiently transformed in tobacco: roGFP2-pmoC, roGFP2-pmoA, pmoC-roGFP2, pmoA-roGFP2 and spBtrunc-roGFP2. Fusion to the N-terminus of spBtrunc was excluded due to the presence of the N-terminal signal peptide. The ratio of roGFP2 excitation at 405 nm to 488 nm gives information regarding the redox state of the protein environment: a low 405/488 ratio indicates a reducing environment such as the cytosol, whereas a higher 405/488 ratio indicates an oxidative environment such as the ER lumen. Expression levels of the pMMO subunits fused to roGFP2 in tobacco were very low, therefore, the experiment was carried out in *Nicotiana benthamiana* with co-expression of p19 in order to obtain the highest possible signal.

The roGFP2 protein on its own, which localizes to the cytosol, and a roGFP2 version containing a signal peptide and the HDEL retention signal (roGFP2-HDEL), localizing to the ER lumen, were used as references for the ratio value in each compartment. As expected, roGFP2 was excited with a higher efficiency at 488 nm, whereas roGFP2-HDEL was excited with a relative higher efficiency at 405 nm, as can be seen by confocal images and by the resulting pixel by pixel ratio image (Figure 3.28). RoGFP2-pmoA and pmoA-roGFP2 both had a higher excitation at 488 nm, similar to cytosolic roGFP2 (Figure 3.28). In each case a pixel by pixel ratio was calculated using a redox ratio analysis software (Fricker, 2016). The mean ratio for cytosolic roGFP2 was 0.33 ± 0.01 (n=10), and for roGFP2-HDEL 0.94 ± 0.02 (n=10). RoGFP2-pmoA gave a value for the ratio of 0.33 ± 0.02 (n=8), and pmoA-roGFP2 gave a value of 0.28 ± 0.02 (n=10), suggesting that the N- and C- termini of pmoA were located in the cytoplasm. Both N- and C- terminal fusions of roGFP2 to pmoC induced disruption of the ER in all cells, and spBtrunc-roGFP2 labeled many bright punctate structures (Figure 3.29), therefore quantitative analysis was not carried out in these constructs as the redox ratio in very

bright regions could be altered and the topology of proteins in distorted ER structures may not have reflected the true topology.

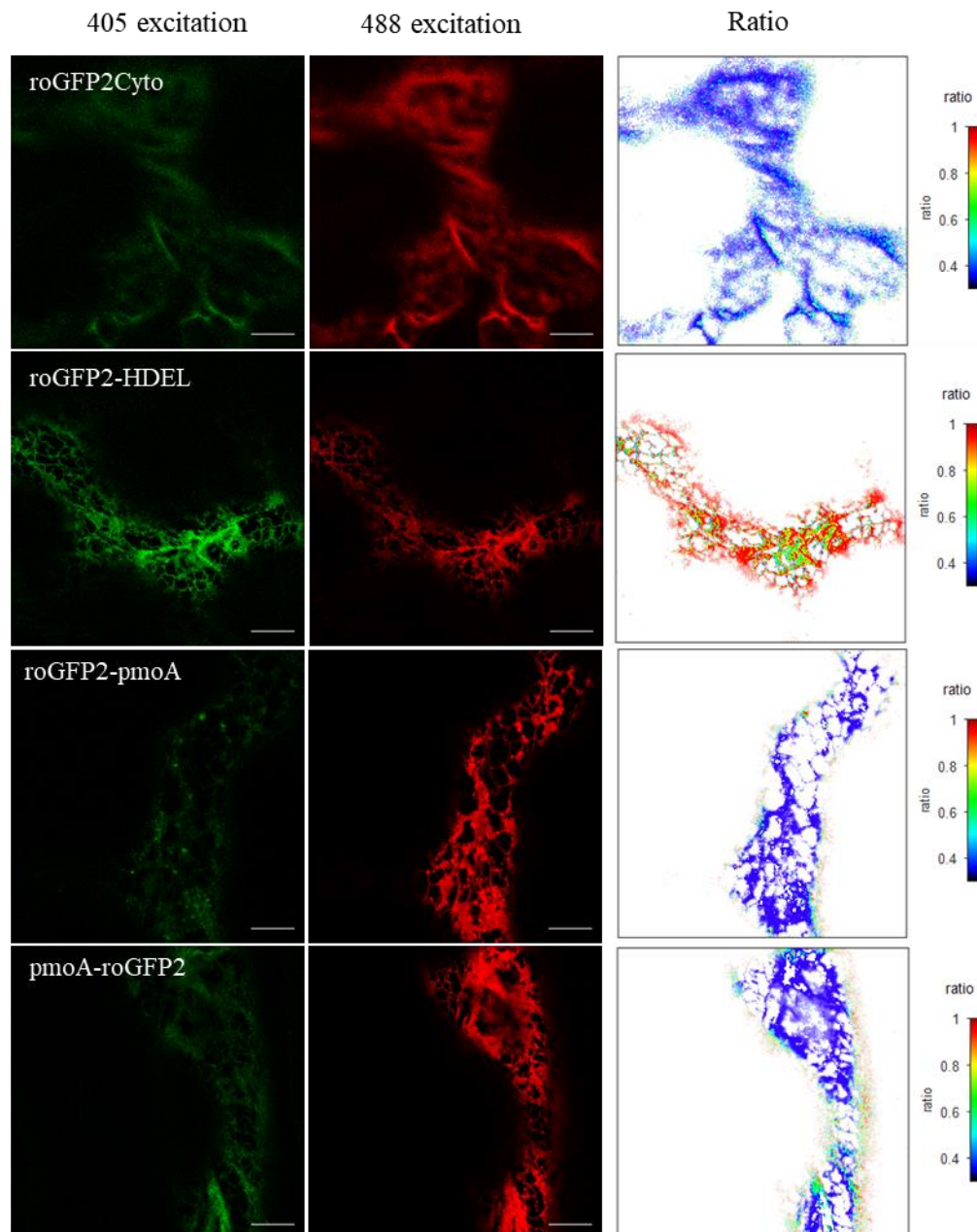


Figure 3.28. Ratio images of roGFP2 constructs.

Confocal images show the emission of roGFP2 fusion proteins with excitation at 405 nm and 488 nm. Cytosolic roGFP2 (roGFP2Cyto), roGFP2-pmoA and pmoA-roGFP2 show a higher signal at 488 nm, and ER luminal roGFP2-HDEL exhibits a comparatively higher excitation at 405 nm. Ratio images show the 405/488 pixel by pixel ratio after processing with redox ratio analysis software. Scale bar = 10 μ m.

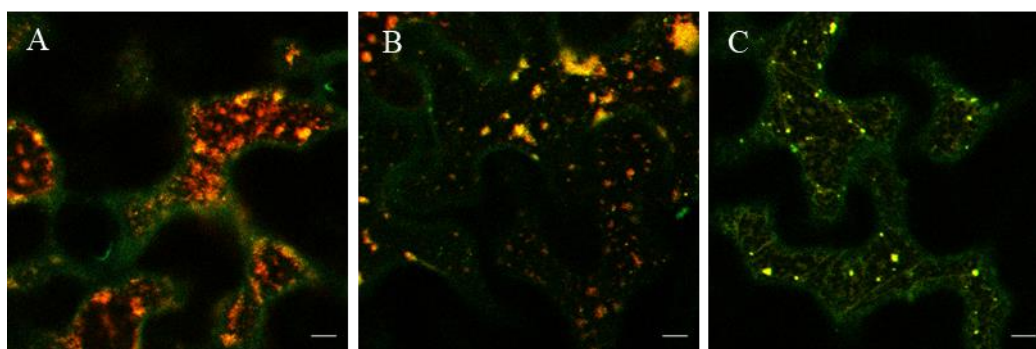


Figure 3.29. Disrupted ER morphology of roGFP2 constructs.

Constructs which were excluded from the ratio analysis due to disruption of the ER network in all cells: (A) roGFP2-pmoC, (B) pmoC-roGFP2 and (C) spBtrunc-roGFP2. Scale bar = 10 μ m.

3.2.6 Co-expression of all three subunits using three-color imaging

With all three bacterial subunits being successfully targeted and localized individually to the ER, the next step in producing the recombinant pMMO complex comprised co-expressing the subunits simultaneously to form the complete complex. In order to analyze the co-expression using microscopy a three-fluorophore approach was used. The pmoA subunit was cloned into a CFP vector in order to co-express with RFP and GFP tagged constructs. CFP-pmoA localized to the ER (Figure 3.30 A and greyscale image Figure 3.30 B showing a clearer contrast). However, most cells expressed with a lower signal to noise ratio than that shown in Figure 3.30 A, partly due to the low quantum yield of CFP and autofluorescence of the cytosol bleeding through into the CFP channel. When other subunits were fused to CFP a clear ER signal could not be detected.

CFP-pmoA was thus co-expressed with spBtrunc-RFP_{hdel} and with GFP-pmoC for three-color imaging, using ODs of 0.1, 0.1, and 0.05 respectively (a lower OD of GFP-pmoC was used to reduce damage to the ER). Most cells showed RFP and GFP signal, with CFP levels indistinguishable from autofluorescence, but also disruption of the ER network (Figure 3.30 C, D, E single channels and F merged channel). However, some ROIs occasionally had cells containing more than one subunit and an intact ER network (merged channel Figure 3.30 G). The presence of pmoA in the ER was very difficult to determine due to the low quantum yield of the fluorophore and autofluorescence causing a diffuse signal throughout the cell. In very

few cases, a cell could be detected harboring all three subunits (Figure 3.30 H), where the pseudo coloring of CFP was changed to light blue to provide enough contrast with a black background, and the nuclear envelope is shown (inset) in all three channels.

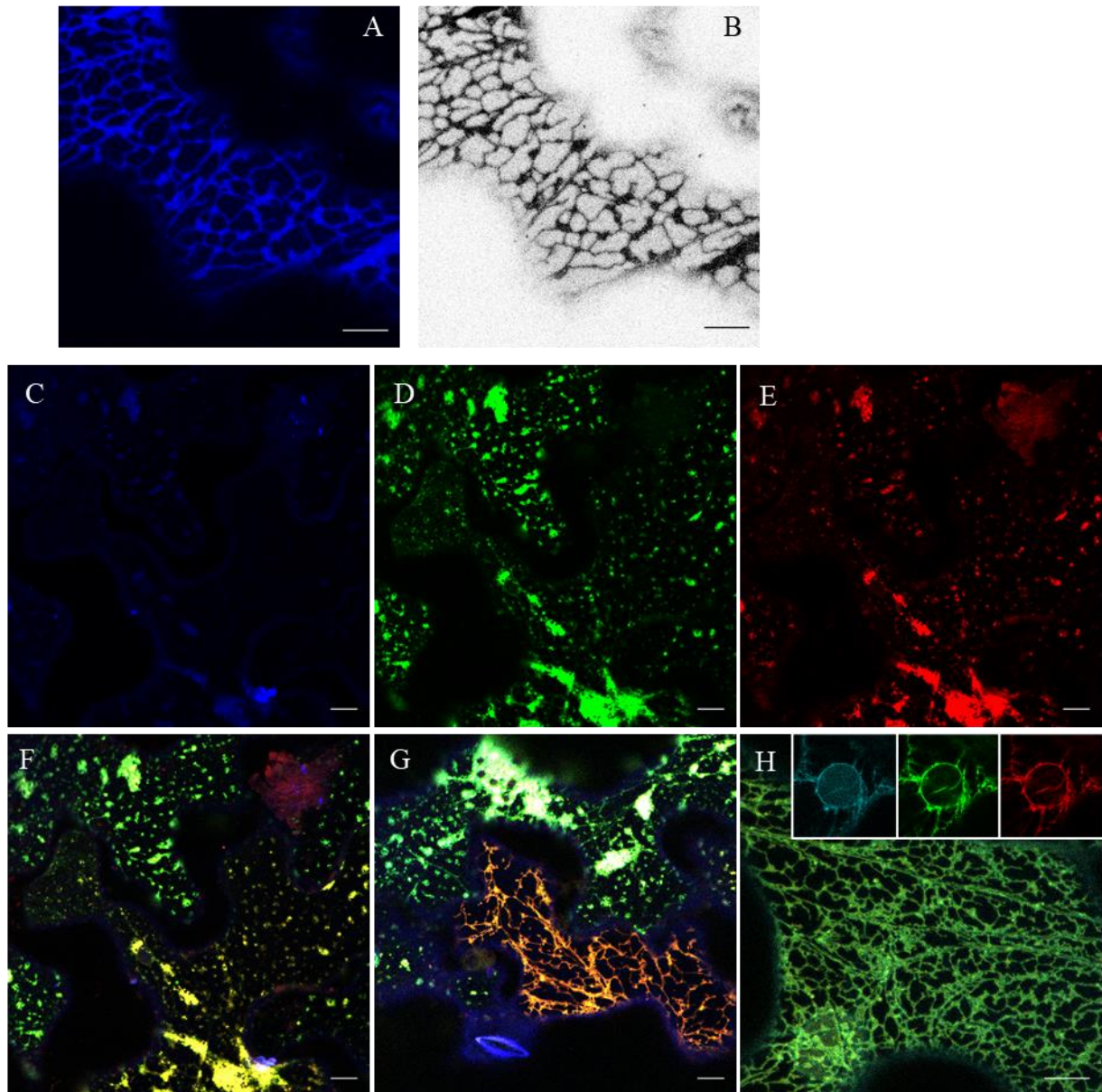


Figure 3.30. Co-expression of pmoC, pmoA and pmoB.

(A) CFP-pmoA expressed in the ER, (B) greyscale image of (A), (C-E) single channel images of CFP-pmoA, GFP-pmoC and spBtrunc-RFPphdel, (F) merged image of (C-E), (G) merged image of CFP-pmoA, GFP-pmoC and spBtrunc-RFPphdel, (H) merged image of CFP-pmoA, GFP-pmoC and spBtrunc-RFPphdel in a healthy cell containing all three subunits, inset: expression of all subunits in the nuclear envelope. Scale bar = 10 μ m.

Thus, generally, the co-expression of the pMMO subunits from different vectors did not result in a viable reconstitution strategy due to a very small number of cells where signal from all three subunits could be found. Moreover, this strategy produced a high proportion of cells with a disrupted ER network, very low detectable levels of CFP-pmoA, and varying RFP to GFP stoichiometries as can be seen in representative images showing the merged channels (Figure 3.30 F and G). In very few cases, all three subunits could be observed with no damage to the ER morphology (Figure 3.30 H), this demonstrated that co-expression can result in ER localized subunits. However, alternative methods for the co-expression of the subunits were required to allow a more coordinated production of the proteins and a less challenging route for the production of stable transgenics which does not rely on a triple co-transformation.

3.3 Discussion

3.3.1 Bacterial pmoC and pmoA localize recombinantly to the plant ER

The pMMO complex is comprised of three subunits (pmoC, pmoA and pmoB), which localize and assemble into a nanomer on the inner membrane of methanotroph bacteria, with the soluble regions of pmoB facing the periplasmic space (Lieberman and Rosenzweig, 2005). The subunits have not been expressed in a eukaryotic cell before. Transient expression in tobacco leaf cells showed localization of pmoC and pmoA in the ER (Figures 3.6 and 3.7). A major route for targeting proteins to the endoplasmic reticulum is the co-translational pathway, where the signal recognition particle (SRP) recognizes an N-terminal signal peptide or transmembrane domain as it emerges from the ribosome, and directs the translating ribosome to the ER membrane (Rapoport, 2007). This transport process is highly conserved across the domains of life, with the ER membrane being analogous to the plasma membrane in Gram-positive bacteria and archaea, or the inner membrane in Gram-negative bacteria (Rapoport, 2007). Due to their high content of transmembrane domains, it is likely that pmoC and pmoA are being targeted to the plant ER via the SRP pathway in an analogous way to their targeting to the bacterial inner membrane in their native host.

Plants were able to transiently express pmoA at high levels at 2 days after infiltration without visible damage to the ER (Figure 3.6). On the other hand, whilst cells expressing pmoC at very low levels showed no visible damage to the ER, cells with medium or high expression presented ER disruption, and tissue necrosis of the infiltrated leaf section was seen at 48 h after infiltration (Figure 3.8). The specific reason for the necrosis caused by pmoC is unknown, however, over-expression of plant proteins or recombinant protein expression can often lead to necrosis as transient expression can result in localized high levels of production. As an example, transient expression of the Ebola glycoprotein GP1 in *Nicotiana benthamiana* caused severe necrosis in the infiltrated leaf sections (Bhoo *et al.*, 2011). Another study reported necrosis of tobacco leaf tissue following transient expression of diverse blood-clot dissolving

proteinases (Yao *et al.*, 2015). Similarly, (Dickey *et al.*, 2017) reported necrosis in tobacco after transiently expressing the earthworm anticoagulant enzyme lumbrokinase, and suggested that the cell damage could have been caused by the hydrolytic activity of the enzyme potentially acting upon native proteins. However, many times necrotic phenotypes are not reported. It could be speculated in the case of pmoC that the phenotype is due to the protein's membrane folding causing destabilization and disruption of the ER membrane.

3.3.2 Bacterial pmoB can be targeted to the plant ER

Recombinant expression of the pmoB subunit from *Methylococcus capsulatus* (Bath) has been previously achieved in *E.coli* using a copper tolerant strain (Lu *et al.*, 2019). The full-length pmoB as well as N-terminal truncations of the protein were reported to localize to the plasma membrane (Lu *et al.*, 2019). Transient expression of pmoB in tobacco plants resulted in localization of the subunit in mitochondria when fused to a C-terminal tag and in the cytosol when fused to an N-terminal tag (Figure 3.10). Therefore the pmoB protein was not recognized by the plant ER targeting machinery in spite of containing a bacterial signal peptide. One study found that a bacterial signal peptide could target proteins to the plant ER (Moeller *et al.*, 2009). Signal peptides are comprised of three regions: a positively charged N-terminal region (“N-region”), a central hydrophobic region (“H-region”), and a C-terminal region which contains the cleavage site (“C-region”) (von Heijne, 1985). When comparing the pmoB signal peptide (from *Methylosinus trichosporium* OB3b) with two different plant signal peptides and with the signal peptide from Moeller et al. (*E. coli* LT-B) which functioned properly in the plant, it can be noted that the pmoB sequence contains a longer N-region and C-region (Figure 3.31). When pmoB is entered into SignalP-5.0 the sequence is not predicted to contain a signal peptide in Eukarya (Figure 3.31 A). When shortening the N-region of the pmoB signal peptide the probability for a signal peptide in Eukarya reaches 1, however, there is no clear probability for a cleavage site (Figure 3.31 B). When shortening both the N-region and the C-region, the sequence is now predicted to contain a signal peptide with a cleavage site in Eukarya in the

correct position (Figure 3.31 C). It is therefore possible that the pmoB signal peptide is not targeting the subunit to the ER in the plant due to its atypically long N and C-regions. This theory could be tested by designing cropped versions of the signal peptide based on *in silico* predictions.

Pumpkin 2S albumin	MA ~~RL~~ TSIIALFAVALLV ~~AD~~ AYAYRT
<i>N. tabacum</i> PR1a	MGFVLFSQLPSFLLVSTLLLFLV ~~ISH~~ SCRA
<i>E. coli</i> LT-B	MV ~~KVK~~ CYVLFTALLSSL ~~C~~ AYG
PmoB	MKAL ~~ER~~ MA ~~EL~~ ATGRV ~~GK~~ LLGLSVAAAVAA ~~TAAS~~ VAPAEA

A. PmoB

B. PmoB with shortened N-region:

~~VG~~KLLGLSVAAAVAA~~TAAS~~VAPAEA

C. PmoB with shortened N and C-regions:

~~VG~~KLLGLSVAAAVAA~~VAP~~AEA

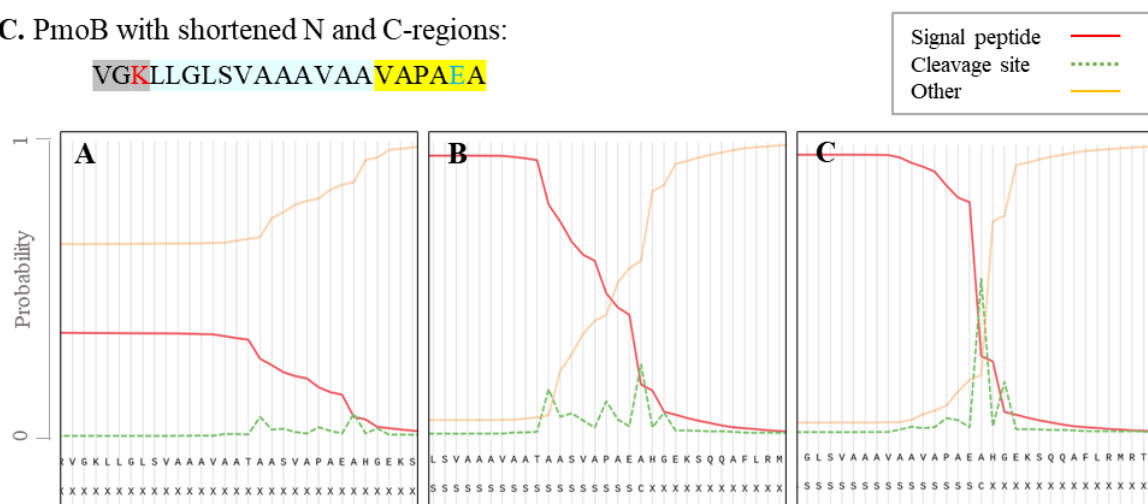


Figure 3.31. Comparison of signal peptide structures.

The different regions of four signal peptides are shown: N-region (grey), H-region (light blue), C-region (yellow). Positively charged residues are marked in red and negatively charged residues are marked in blue. The SignalP-5.0 prediction for a signal peptide in Eukarya is shown for pmoB (A), pmoB with a shortened N-region (B), and pmoB with a shortened C-region (C). The probability for a signal peptide across the protein sequence is marked as a red line, the probability for a cleavage site in a green dotted line and the probability of no signal peptide is marked as a yellow line.

Nevertheless, changing the pmoB signal peptide to a plant signal peptide resulted in an appropriate strategy to target pmoB to the ER (Figure 3.16). Expression levels of the

recombinant protein were considerably lower than those of the other subunits. Fusion proteins of the ER targeted pmoB (spBtrunc) to RFP showed that some of the protein was localized to the vacuole after 2 days and entirely to the vacuole after 3 days, suggesting that the protein was not correctly retained in the ER. An HDEL retention signal was effective at retaining spBtrunc-RFP-HDEL in the ER both at 2 days and 3 days after infiltration (a summary of all constructs related to targeting is found in Appendix 3 – section 9.2). Surprisingly however, expression levels of spBtrunc-GFP-HDEL were only fractionally higher than spBtrunc-GFP (Figure 3.20). However, upon co-expression with the silencing suppressor P19, spBtrunc-GFP-HDEL levels were markedly increased (Figure 3.21). The PTGS suppressor P19 has frequently been used to enhance recombinant protein production in plants (Peyret and Lomonossoff, 2015). In this case, the expression of spBtrunc-RFP-HDEL or the co-expression of spBtrunc-GFP-HDEL with P19 resulted in two ways of achieving high levels of pmoB in the plant ER, with no visible alterations to ER morphology. It remains unclear whether and why a PTGS mechanism may be acting to a higher extent in the GFP construct.

3.3.3 Recombinant pMMO proteins interact with each other

In order to achieve a functional version of pMMO, the subunits must be correctly oriented with respect to one another and native interactions have to be reconstituted in the plant. The only previous report of recombinant expression of pMMO showed very low levels of activity compared with methanotroph cells (Gou *et al.*, 2006). This could have been due to many reasons, and one possible question stated by the authors is whether the complex had assembled correctly in their system. Here, pairwise protein-protein interactions were assessed with FRET-FLIM. Likely interactions based on a significant reduction in donor lifetime were found between both termini of pmoA and those of pmoC (Figure 3.25). The reductions in lifetime ranged between 0.13 and 0.26 ns, which partially overlaps with decreases previously reported for interacting partners in other studies, which were typically greater or equal to 0.2 ns (Stubbs *et al.*, 2005; Kriechbaumer *et al.*, 2015, 2016). The distance between fluorophores and

alignment of the transition moments are two factors affecting FRET efficiency. It is possible that the slightly lower decreases found here could be due to a lower FRET efficiency linked to the protein conformations altering the distance and alignment of the fluorophores.

Distances between the termini in the crystal structure provide a guide on expected interactions, even though it is not guaranteed that if expressed pairwise the subunits would assemble identically as in the trimer structure. The N-termini of pmoC and pmoA are highly proximal in the crystal structure (Figure 3.32 D), therefore it was expected for GFP-pmoA and mCherry-pmoC to interact. However, the C-terminus of pmoA in the crystal structure is located on the other side of the membrane, and sits at a distance of almost 7 nm from the N-terminus of pmoC (Figure 3.32 C), therefore, considering also the diameter of the fluorophores, it was not expected that pmoA-GFP and mCherry-pmoC would interact. The C-terminus of pmoC is not modelled in the crystal structure, therefore a distance was not calculated in PyMOL, however, it is expected to lie on the same side of the membrane as the N-terminus of pmoC. The distance between the pmoB C-terminus and pmoA C-terminus is also just below 7 nm (Figure 3.32 A), therefore an interaction seemed unlikely, however, being on the same side of the membrane this could have depended on how the fluorophores were positioned with respect to each other. Finally, the separation between the C-terminus of pmoB and N-terminus of pmoC was also too large to expect an interaction (Figure 3.32 B), and no interaction was observed in this case. The expected interactions and those found via FRET-FLIM are summarized in Table 3.4. Due to the strong prediction of both pmoC termini facing the cytosol, the most plausible explanation for an interaction between pmoA-GFP and both pmoC termini is that the C-terminus of pmoA might reside in the cytosol instead of in the ER lumen.

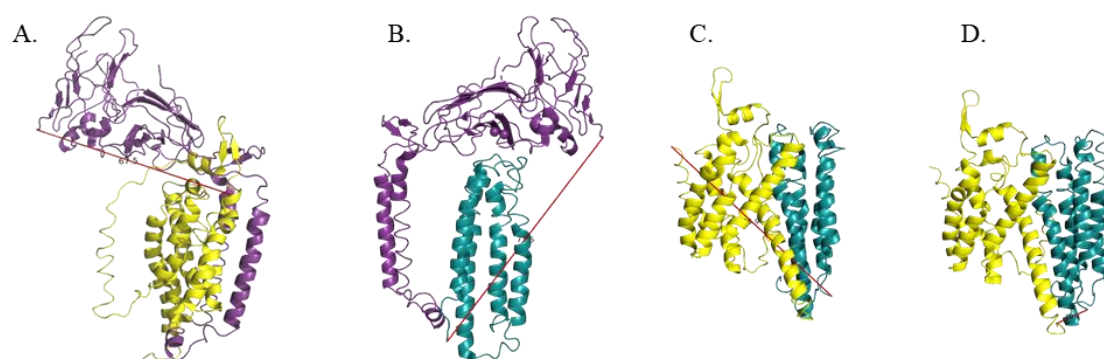


Figure 3.32. Distance between protein termini calculated in PyMOL.

Cartoon representation of pmoB (purple), pmoA (yellow), and pmoC (turquoise), showing (A) distance between the C-terminus of pmoB and C-terminus of pmoA (67.5 Å), (B) distance between the C-terminus of pmoB and the N-terminus of pmoC (73.5 Å), (C) distance between the N-terminus of pmoC and C-terminus of pmoA (65.9 Å), (D) distance between the N-terminus of pmoC and N-terminus of pmoA (10.2 Å). The C-terminus of pmoC was not included due to a gap in the protein structure. Distances are shown with a red line.

Table 3.4. Expected and found interactions with FRET-FLIM.

Expected interactions based on distances calculated from the crystal structure and predicted membrane topology, and interactions found via FRET-FLIM.

Donor	Acceptor	Expected interaction	Found interaction
N-terminus pmoA	N-terminus pmoC	Yes	Yes
N-terminus pmoA	C-terminus pmoC	Yes	Yes
N-terminus pmoA	C-terminus pmoB	No	No
C-terminus pmoA	N-terminus pmoC	No	Yes
C-terminus pmoA	C-terminus pmoC	No	Yes
C-terminus pmoA	C-terminus pmoB	Possibly	No
N-terminus pmoC	C-terminus pmoB	No	No

Other methods to test protein-protein interactions such as co-immunoprecipitation (co-IP) could complement FRET-FLIM data to confirm interactions between subunits, especially between pmoB and pmoA where FRET-FLIM did not detect an interaction, potentially because of the fluorophores being on different sides of the membrane. An interaction between pmoB and pmoA is likely due to retention of spBtrunc-RFP in the ER observed when both subunits are co-expressed, as opposed to the fully vacuolar localization of spBtrunc-RFP without pmoA (Figure 3.27). Furthermore, if a method was obtained to express all three subunits simultaneously and at high levels, protein extraction in non-denaturing conditions followed by

native PAGE gels could be used to assess the assembly of the complex in the trimer or nanomer form.

3.3.4 Topology of pmoA

Protein topology was assessed using the roGFP2 system. Fusions of roGFP2 to pmoC resulted in high levels of ER disruption and therefore pmoC was excluded from the analysis. However, predictions of pmoC topology show a high consensus for both termini to reside in the cytoplasm. The fusion of roGFP2 to pmoB resulted in the formation of bright puncta which were not found in the GFP tagged version of this construct, and which could likely alter the 405/488 nm ratio, thus the construct was also not analyzed with the redox ratio software. However, due to the presence of a signal peptide and effective retention in the ER lumen with the HDEL motif, it can be deduced that the termini of pmoB reside in the ER lumen. Finally, the redox analysis of pmoA showed that both termini of pmoA reside in the cytosol, with a ratio of 405/488 around 0.3 similar to that of the cytosolic control. This topology differs from the folding in the crystal structure which shows the C-terminus of pmoA residing in the periplasmic space of bacteria, which would translate to the lumen of the ER. However, the topology found via the redox ratio analysis does coincide with the protein-protein interactions found with FRET-FLIM, where the C-terminus of pmoA interacts with both termini of pmoC, assumed to be in the cytosol.

Interestingly, despite the modified topology, pmoA seems to still be able to interact with the other subunits. When comparing the pmoA structure used in this study (PDB 3CHX) with the crystal structures available from other genera of methanotrophs it is observed that all other versions of pmoA have both termini facing the cytosol (Figure 3.33). Moreover, all structures seem relatively conserved, except for the loop region after a small alpha helix (orange, Figure 3.33): the 3CHX structure shows the loop traversing the membrane (indicated with a white arrow) and then a final loop partially modelled as an alpha helix traversing the membrane once more yielding the topology with the periplasmic C-terminus, whereas all other structures (B-

D) show the loop modelled parallel to the membrane on the periplasmic space, and the final alpha helix traversing the membrane yielding a topology with a cytosolic C-terminus. The amino acid sequence of the loop is highly conserved between the species compared (Figure 3.34), therefore the primary sequence in itself would not support a large difference in the structure, although it cannot be ruled out that interactions between the residues in the loop and other components of the structure could cause this difference in folding. Furthermore, phylogenetic distances between the genera would also not explain a priori the groupings in the structures as *Methylosinus* (3CHX) and *Methylocystis* (3RFR) belong to the same family (Methylocystaceae), and *Methylococcus* (3RGB) and *Methylomicrobium* (6CXH) belong to the family Methylococcaceae.

One possibility is that the 3CHX structure could harbour an error in the loop region. Crystal structures are highly reliable however they may contain errors associated with protein disorder, such as in flexible loop regions, and lower accuracy as the resolution decreases (Mackay *et al.*, 2017). The resolution of 3CHX is relatively low (3.9 Å), and is also lower than the resolution of the other pMMO structures (around 2.7 Å). In addition, the structure validation report of pmoA in 3CHX shows a high number of residues containing outliers in several geometric quality criteria, whereas 3RFR shows very few residues containing outliers (Supplementary Figure 3.3). Therefore, one possibility is that the native folding of the loop in pmoA is the one shown in structures B-D and the folding of structure A may have errors. In this case, the topology of pmoA found in this study would agree with that of the crystal structure. Another possibility, although less likely, is that pmoA from *Methylosinus trichosporium* folds differently to pmoA from the other characterized methanotrophs. In this case, pmoA folding in plants would differ from the folding in the bacteria, although still allowing for interactions with other subunits and therefore potentially still enabling the assembly of a functional enzyme.

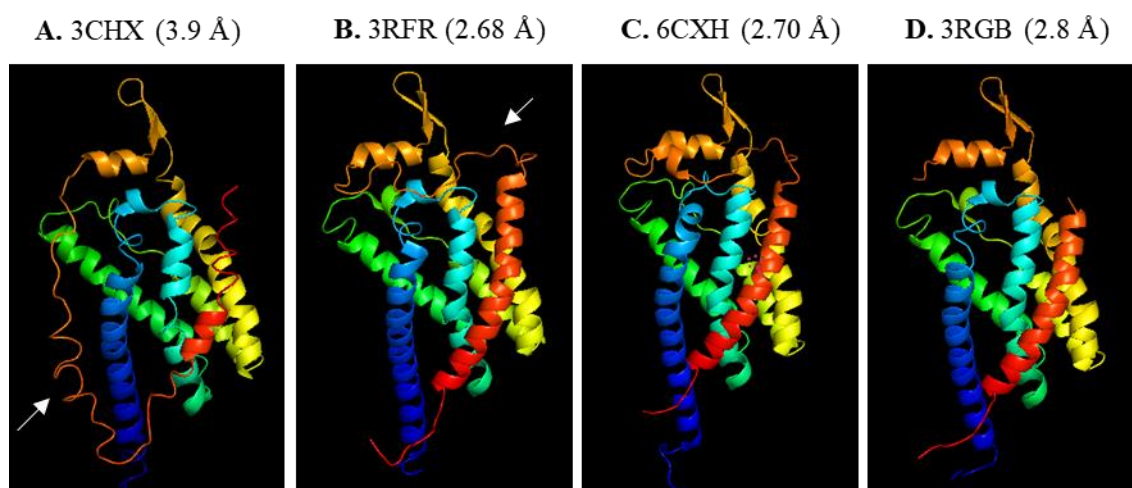


Figure 3.33. Folding of pmoA in different methanotroph genera.

Crystal structures of the pmoA subunit from (A) *Methylosynus trichosporium* (3CHX), (B) *Methylocystis* sp. strain M (3RFR), (C) *Methylomicrobium alcaliphilum* 20Z (6CXH), and (D) *Methylococcus capsulatus* (Bath) (3RGB). The resolution of the full pMMO structure in each case is shown in brackets. The loop which differs between 3CHX and the remaining structures is indicated with a white arrow, and is missing in the 3RGB structure due to unmodelled residues. The N-terminus (beginning of the blue helix) is facing the cytosol in all four structures, whereas the C-terminus (end of the red helix) is facing the bacterial periplasm in 3CHX and the cytosol in the remaining three structures.

CLUSTAL O(1.2.4) multiple sequence alignment

```

3CHX      MFTSKSGGAIGPFHSVAEAAAGCVKTTDWMFLTLLFLAVLGGYHIHFMLTAGDWDFFWVDWK 60
3RFR      MSQSKSGGAVGPFNSVAEAAAGCVATTDWMLLVLLFFAVLGGYHVFHMLTAGDWDFFWVDWK 60
6CXH      -----MSASQSAVRSRAEAVKVSRTFDYMILFTVFFVVLGGYHIHYMLTGGDWDFFWTDWK 55
3RGB      -----MSAAQSAVRSRAEAVQVSRITIDWMLFVVFVIVGSYHIHMLTMDGWDFFWSDWK 55
          .:. . . * * * . * * * * : * . . . * * * * * * * * * * * *

3CHX      DRRMWPTVVPILGVTFAAAAQAFFWENFKLPFGATFAVSGLLIGEWINRYCNFWGWTYFP 120
3RFR      DRRMWPTVLPILGVTFCASQAFWVNFRLPFGAVFAVLGLMIGEWINRYVNFWGWTYFP 120
6CXH      DRRLWTVAPIVSITFPAAVQAVLWRYRIAGATLCVLGLLGEWINRYFNFWGWTYFP 115
3RGB      DRRLWTVTPIVLVTFAAVQSYLWERYRLPWGATVCVLGLLGEWINRYFNFWGWTYFP 115
          * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

3CHX      ISLVFPSALVVPALWLDIIMLLSGSYVITAVVGSGLWGLLFYPPNNWPAIAALHQATEQHG 180
3RFR      ISLVFPSAMIVPAIWLVDIILLSGSYVITAVVGSGLWGLLFYPPNNWPAIAAFHQATEQHG 180
6CXH      VNFVFPNLMPGAIVLDVILMLSNSMTLTAVVGGGLWGLLFYPPGNWPIIAPLHVPVEYNG 175
3RGB      INFVFPASLVPGAIIIDTVLMLSGSYLFTAIVGAMGWGLIFYPPGNWPIIAPLHVPVEYNG 175
          .:. * * * : : * : * * : : * * * * * * * * * * * * * * * * * *

3CHX      QLMSLADLVGFHFVRTSMPEYIRMVERGTLRTFGKEVVPVAAFFSGFVSMMVYFLWVFVG 240
3RFR      QLMTLADLIGLHFVRTSMPEYIRMVERGTLRTFGKDVVPVAAFFSGFVSMMVYFLWVFVG 240
6CXH      MMTLADLQGYHYVRTGTPEYIRMVEKGTLRFTFGKDVPVSAFFSGFVSILYFLWHFFG 235
3RGB      MLMSIADIQGYNYVRTGTPEYIRMVEKGTLRFTFGKDVPVSAFFSAFMSILYFLWHFFG 235
          : : * * * * * * * * * * * * * * * * * * * * * * * * * * * *

3CHX      KWYSTTKVIQKI 252
3RFR      RWYSTTKRIEQI 252
6CXH      SWFGSEKVFQAA 247
3RGB      RWFNSNERFLQST 247
          * : . . : : :

```

Figure 3.34. Alignment of the protein sequence of the pmoA subunits from different methanotroph genera.

Clustal Omega alignment with coloring according to residue type. The amino acids which form the loop region which differs between 3CHX and the other structures are marked by a rectangle and show high sequence conservation.

4 Strategies for gene stacking

4.1 Introduction

4.1.1 Gene stacking with multiple promoters

Biotechnology in plants increasingly requires the incorporation of multiple transgenes, known as “gene stacking” (Halpin, 2005), to introduce novel traits and to modify metabolic pathways. Multi-gene transformation of plants is more complex than in bacteria as the latter can be transformed with multiple genes in an operon, whereas in eukaryotes each gene requires its own promoter and terminator sequences and has historically relied on multiple independent transformation events, and crossing transgenic lines each harbouring a transgene. For example, two Broccoli lines containing different *Bacillus thuringiensis* (Bt) genes which confer protection against diamondback moths were crossed, and resulted in lines which caused mortality of moths resistant to one protein or the other (Cao *et al.*, 2002). One drawback of this method is that it is very time consuming and labour intensive due to the long life cycles of plants and the requirement to grow and characterise several generations to obtain homozygous plants, exacerbated by doing this sequentially for the addition of further genes. Other drawbacks comprise the need for multiple selectable marker genes, which could result in approval problems for implementing modified varieties in the field, and segregation events due to transgenes incorporated in different genomic loci, which could lead to loss of a trait in subsequent generations (Halpin, 2005). This is thus not an ideal method, especially if requiring the simultaneous introduction of more than two genes.

Another approach to gene stacking is co-transformation, where genes can be incorporated simultaneously in the same transformation event, either from separate plasmids, or from one vector containing multiple expression cassettes (Figure 4.1 A and B). Transforming with separate plasmids still requires multiple selectable markers, with a low probability of getting a

stable transformant, and in transient expression the stoichiometry between genes can be very variable. Instead, using a plasmid which can harbour multiple cloning cassettes or transcriptional units would circumvent these issues. For example, “2in1” vectors contain multiple cassettes that are compatible with Gateway cloning technology within a single T-DNA region for agrobacterium-mediated transformation (Grefen and Blatt, 2012). A remarkable case of co-transformation was the introduction of the provitamin A (beta-carotene) biosynthetic pathway into rice endosperm, for the creation of “golden rice”, designed to combat vitamin A deficiency in under-developed countries (Ye *et al.*, 2000). In this work the authors employed two different T-DNAs comprising two genes each, which were introduced simultaneously by Agrobacterium-mediated transformation (Ye *et al.*, 2000). An increasingly popular approach for gene stacking is the use of the Modular Cloning (MoClo) system which is based on the Golden Gate cloning method, where multiple fragments can be assembled sequentially through the use of Type IIS restriction enzymes, creating tandem transcription units within a single vector (Engler *et al.*, 2008; Weber *et al.*, 2011).

The advantages of co-transformation over crossing or re-transformation include a reduction in the time and labour, and, when using linked genes in a same T-DNA, a reduction in the number of marker genes needed, avoidance of segregation, and more homogeneous expression between cells. A limitation which arises from the method is still a possible lack of coordination in protein expression due to the use of separate promoters, and also the possibility of downregulation or silencing with the use of multiple promoters containing homologous regions (Rajeev Kumar *et al.*, 2015). Thus a range of strategies based on polycistronic vectors (or polypeptides) have been developed to address these drawbacks.

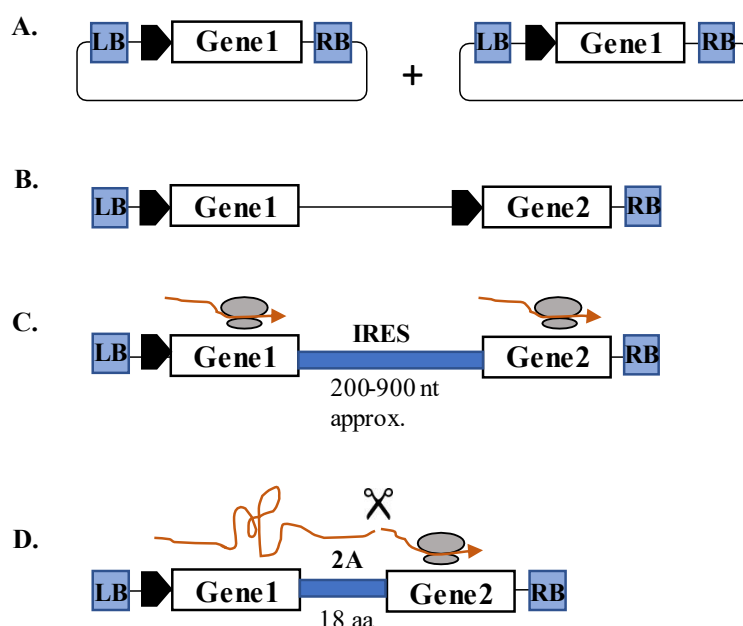


Figure 4.1. Methods for gene stacking.

Schematic diagram showing different co-expression techniques: co-expression from separate plasmids (A), co-expression within a same T-DNA region (B), polypeptide mediated by IRES (C), polypeptide mediated by self-cleaving peptide (D). The proteins being synthesized are shown in orange for (C) and (D). LB= left border, RB= right border, black arrow= promoter.

4.1.2 Gene stacking with a single promoter: polypeptides

Polypeptides are constructs which contain several genes under a same promoter that are compatible with the eukaryotic translation machinery. One method to create polypeptides uses internal ribosome entry sites (IRES) to separate genes (Figure 4.1 C). IRES sequences are present in the 5'UTR region of some genes and mediate a cap independent initiation of translation (Urwin *et al.*, 2000). Originally identified in viruses, they have subsequently been found in other cell types, and are typically present in genes which contain an extended 5'UTR region (Urwin *et al.*, 2000). IRES vary in length and neither their amino acid sequence nor their secondary structure are conserved between organisms, so their activity has to be confirmed empirically (Ngoi *et al.*, 2004). Synthetic polypeptides containing IRES sequences have been used in plants and shown to mediate the translation of two genes (Urwin *et al.*, 2000; Gouiaa *et al.*, 2012). However, these sequences have the drawbacks of being fairly long (ranging from 130 bp to 1kb, but typically around 500 bp or more), and of having a relatively low translational

efficiency. Notably the downstream gene has often been reported to translate at levels of around 10% of the upstream gene (Halpin, 2005; Minskaia and Ryan, 2013). These issues are particularly relevant when expressing more than two genes, and when the stoichiometry of the different components is important.

Another existing methodology to construct polypeptides uses “2A” self-cleaving peptides (Figure 4.1 D). The first 2A peptide to be discovered derived from the foot-and-mouth disease virus (FMDV), termed F2A, and is a short 18 amino acid peptide that produces cleavage of polyproteins by a ribosomal skipping mechanism during translation between the 17th and 18th amino acid of the peptide (Donnelly *et al.*, 2001). This process leaves the upstream protein with an addition of 19 amino acids on its C-terminus, whereas the downstream protein incorporates the 20th amino acid - a proline residue - on its N-terminus (Donnelly *et al.*, 2001). 2A peptides from other viruses have also been discovered, such as E2A (from equine rhinitis virus), P2A (from porcine teschovirus -1) and T2A (*Thosea asigna* virus), which all contain the conserved C-terminal sequence -GDVExNPGP.

2A peptides have a high cleavage efficiency and have been used successfully for polyprotein construction in a variety of systems including mammalian, insect, plant, fungal and yeast cells, for biomedical and biotechnological purposes (Luke *et al.*, 2015). It has been widely reported that this system is more effective and efficient for co-expression of genes than the use of IRES, due to their shorter length, higher translational levels of downstream proteins, functionality across different cell types, and ability to concatenate more than two transgenes (Halpin, 2005; Ha *et al.*, 2010; Hee Kim *et al.*, 2011; Minskaia and Ryan, 2013; Luke *et al.*, 2015). As an example of their advantages in plant systems, one study co-expressed the carotenoid biosynthetic genes (*Psy* and *CrtI*) both with 2A and IRES, and found that 2A is more efficient at coordinating gene expression and driving translation (Ha *et al.*, 2010). Another study showed the benefit of 2A linked genes over other co-expression techniques in obtaining ratiometric quantities of proteins, which were used to develop secretion assays (Samalova *et al.*, 2006).

Finally, this system has proved beneficial in assembling multicomponent protein complexes. For example, Lin and co-workers found a higher yield of Ebola monoclonal antibodies produced in tobacco by co-expressing the heavy and light chains with a 2A peptide, compared to expressing them from separate transcription units (Lin *et al.*, 2018). They found similar expression levels of the heavy and light chains which they hypothesized could be the cause of the higher yield of the assembled complex. It would be expected for proteins in a polypeptide to be produced at a near 1:1 stoichiometry. However, this is not always the case, as protein quantities can be affected by other processes such as proteasomal degradation, ribosome drop-off along the length of the transcript, ribosome fall-out after the first self-cleaving peptide (Liu *et al.*, 2017), the gene order (Wang *et al.*, 2015), the order and type of self-cleaving peptides (Liu *et al.*, 2017), and the nature of protein additions adjacent to self-cleaving peptides (Reinhardt *et al.*, 2020), amongst others. Nevertheless, regardless of the stoichiometry achieved, this method confers much higher coordination of gene expression compared to other co-expression strategies, due to protein synthesis from a single mRNA molecule, leading to improved ratiometric protein production (Samalova *et al.*, 2006).

Two draw-backs in the 2A method are that a small number of amino acids are added to the C-terminus of the upstream protein, and a proline residue is added to the N-terminus of the downstream protein, which could potentially interfere with protein function or folding, and that cleavage efficiency can vary depending on the construct, leading to uncleaved products which can lower the productivity of the system.

4.1.3 Objectives

In Chapter 3 it was shown that pmoC and pmoA are targeted to the ER in plants and that pmoB can be targeted to the ER by replacing the bacterial signal peptide with a plant signal peptide. However, in order to express all three subunits simultaneously, a triple co-infiltration with the subunits in separate plasmids was attempted but did not result in a reliable approach to give stoichiometric expression. The aim of this chapter was to find an alternative gene stacking

method which would allow successful co-expression of all three subunits. The objectives were to test different gene stacking methods and evaluate:

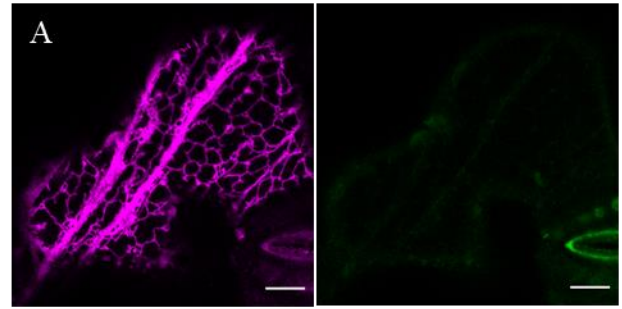
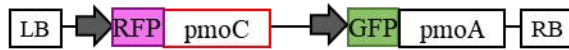
- i) The ability to co-express two or more subunits with detectable expression levels
- ii) Correct ER localization of the subunits
- iii) The presence of protein-protein interactions
- iv) Expression of the subunits in ratiometric quantities.

4.2 Results

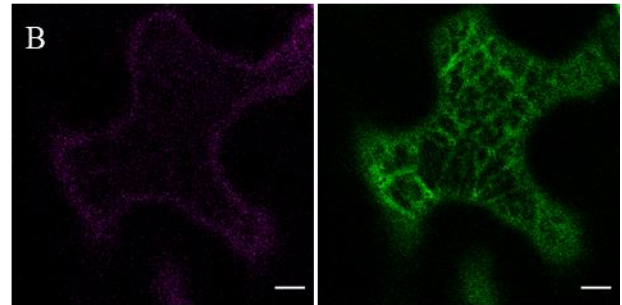
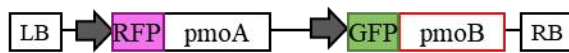
4.2.1 Combinatorial design of “2in1” vectors for protein co-expression

The first approach to achieve gene stacking of the pMMO subunits was to use binary plant 2in1 vectors (Hecker *et al.*, 2015). These vectors contain two cloning cassettes within the same T-DNA region. Each cassette includes a strong constitutive promoter (35S) and a fluorescent protein in frame to allow the fusion of the gene of interest. The method thus enables the simultaneous transformation of two proteins and was reported to increase the frequency of co-transformation in comparison to co-transformation with separate plasmids (Hecker *et al.*, 2015). In this study 2in1 vectors were constructed using different combinations of pMMO subunits, cassette position (first or second cassette), and position of the fluorescent tag (N- or C- terminal), to achieve co-transformation of two subunits. Initially four vectors were constructed to test expression of the cassettes as shown in Figure 4.2 A-D. However, the 2in1-1 vector showed expression solely from the first cassette, whereas 2in1-2, 2in1-3 and 2in1-4 showed expression only from the second cassette (Figure 4.2 A-D). Note: the pmoB sequence used for this experiment was the original bacterial sequence without replacing the signal peptide or adding a retention tag as this was conducted in parallel to the targeting studies described in Chapter 3, to test the feasibility of co-expression systems. For this reason expression of GFP-pmoB is cytosolic (see section 3.2.1.3). Given the strong expression of the first cassette of 2in1-1 and the second cassette in 2in1-3 (and 2in1-4), these cassettes were combined to form 2in1-5, in the hope of obtaining dual expression. However, this construct showed no expression of either cassette (Figure 4.2 E, merged channel). A final construct was created with the same fluorophore positions as 2in1-3 but swapping the subunits so that pmoC was in the second cassette, which had previously given strong expression of pmoA. Again, no expression was observed in either channel (Figure 4.2 F, merged channel). Therefore the 2in1 method permitted the expression of a single subunit, but did not produce dual expression in any of the combinations tested.

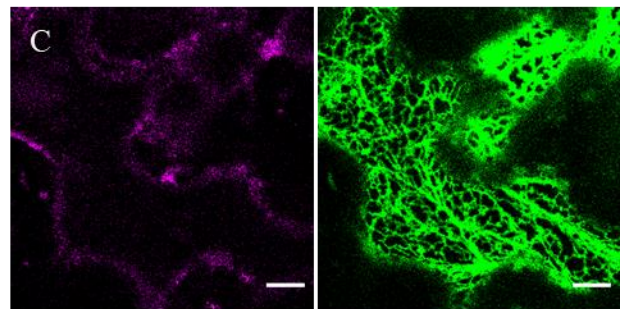
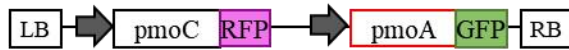
A. 2in1-1



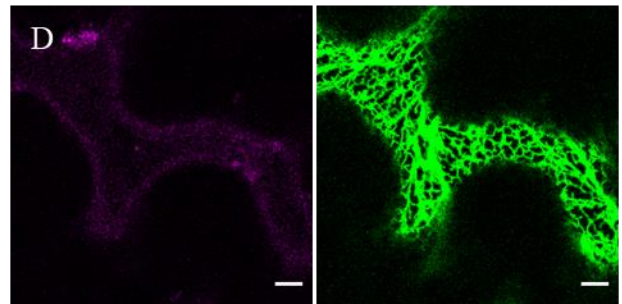
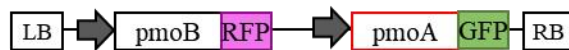
B. 2in1-2



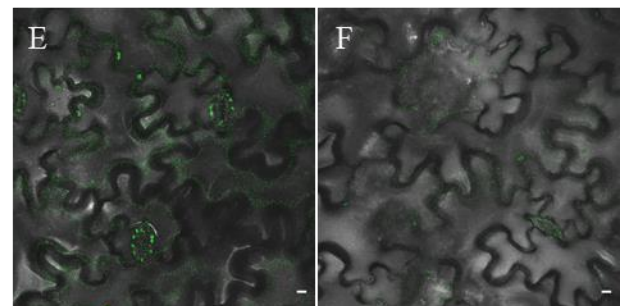
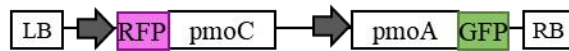
C. 2in1-3



D. 2in1-4



E. 2in1-5



F. 2in1-6

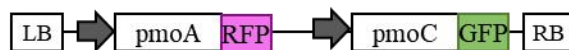


Figure 4.2. Diagram and expression pattern of 2in1 vectors.

(A-D) Simplified diagrams of the 2in1 vectors constructed, accompanied each by a representative image of the expression pattern in tobacco at 2 days after infiltration, showing

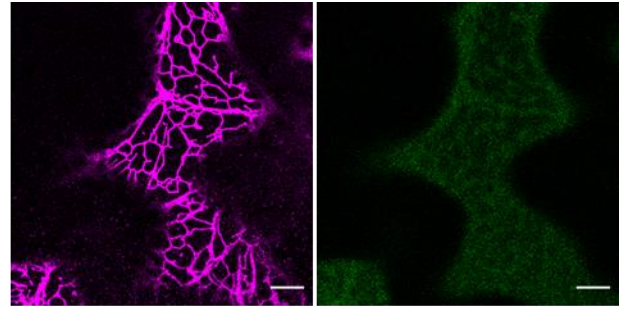
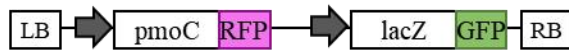
expression from the first cassette (red) and second cassette (green) from the same cell imaged in dual channel mode. Cassettes which showed expression are outlined in red. (E, F) Diagrams of 2in1 vectors which did not show any expression, accompanied by a lower magnification view of the merged channels (red, green and transmitted light). Scale bar = 10 μ m.

4.2.2 Single cassette 2in1 vectors

2in1 vectors failed to produce simultaneous expression of both cassettes with the introduction of pMMO proteins. The lack of expression of 2in1-5, which combined two cassettes which previously showed expression, suggested a possible inhibition acting between the pMMO subunits. To test this, two constructs were created similar to 2in1-3 and 2in1-4, but *pmoA* was replaced in the second cassette by the *lacZ* gene from the vector backbone. These were termed “single cassette” vectors, and used to test if the absence of *pmoA* would lead to expression of the first cassette.

Following transient transformation, the single cassette 2in1-3 construct showed strong ER signal in the red channel (Figure 4.3 A), which was absent in 2in1-3 containing both subunits. Likewise, single cassette 2in1-4, now expressed *pmoB*-RFP, which is seen as bright aggregates with a cytosolic background (Figure 4.3 B). This confirmed that the presence of two pMMO subunits was causing an inhibition in the expression of at least one cassette in 2in1 vectors.

A. Single cassette 2in1-3



B. Single cassette 2in1-4

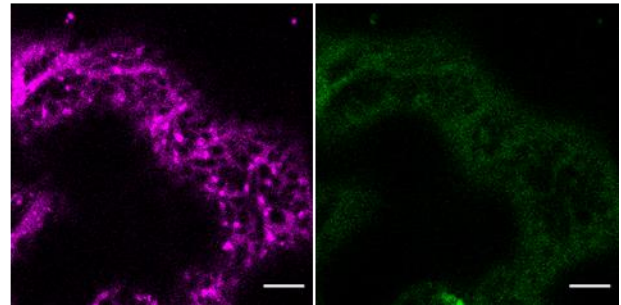
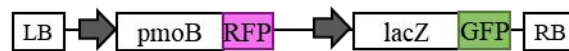


Figure 4.3. Diagram and expression pattern of single cassette 2in1 vectors.

On the left, diagrams show the structure of the 2in1 vector, where only the first cassette underwent recombination to insert a gene of interest. On the right, confocal images of expression after 2 days in tobacco cells. Both constructs express the subunit which could not be seen when both pMMO genes were present (see Figure 4.2 C and D). Scale bar = 10 μ m.

4.2.3 Co-expression of 2in1 vectors with p19

The co-expression of transgenes with p19, a suppressor of post-transcriptional gene silencing (PTGS), has often been reported to boost expression levels. Therefore, in an attempt to obtain improved co-expression with the 2in1 method, the vectors 2in1 1-6 were co-transformed along with the silencing suppressor p19 in *Nicotiana benthamiana*. It was found that 2in1-1 showed expression of the first cassette as expected, but also a very faint expression of the second cassette in *Nicotiana benthamiana*, which was not observed in tobacco (Figure 4.4 A). However, co-expression with p19 did not result in improved expression of the second cassette. 2in1-2 gave a cytosolic signal from pmoB in the second cassette (Figure 4.4 B), whilst 2in1-3 and 4 showed a strong ER signal of pmoA in the second cassette (Figure 4.4 C and D), in a similar manner to what was observed in tobacco. The addition of p19 did not result in improved expression of the silenced cassette in these cases. In the cases of 2in1-5 and 6, where both cassettes appeared to be silenced in tobacco, lack of expression from both cassettes was also

observed in *Nicotiana benthamiana*, with and without the addition of p19 (Figure 4.4 E and F). Thus, the addition of p19 did not produce additional expression in any of the cases, indicating that a different silencing mechanism is taking place in the dual cassette vectors.

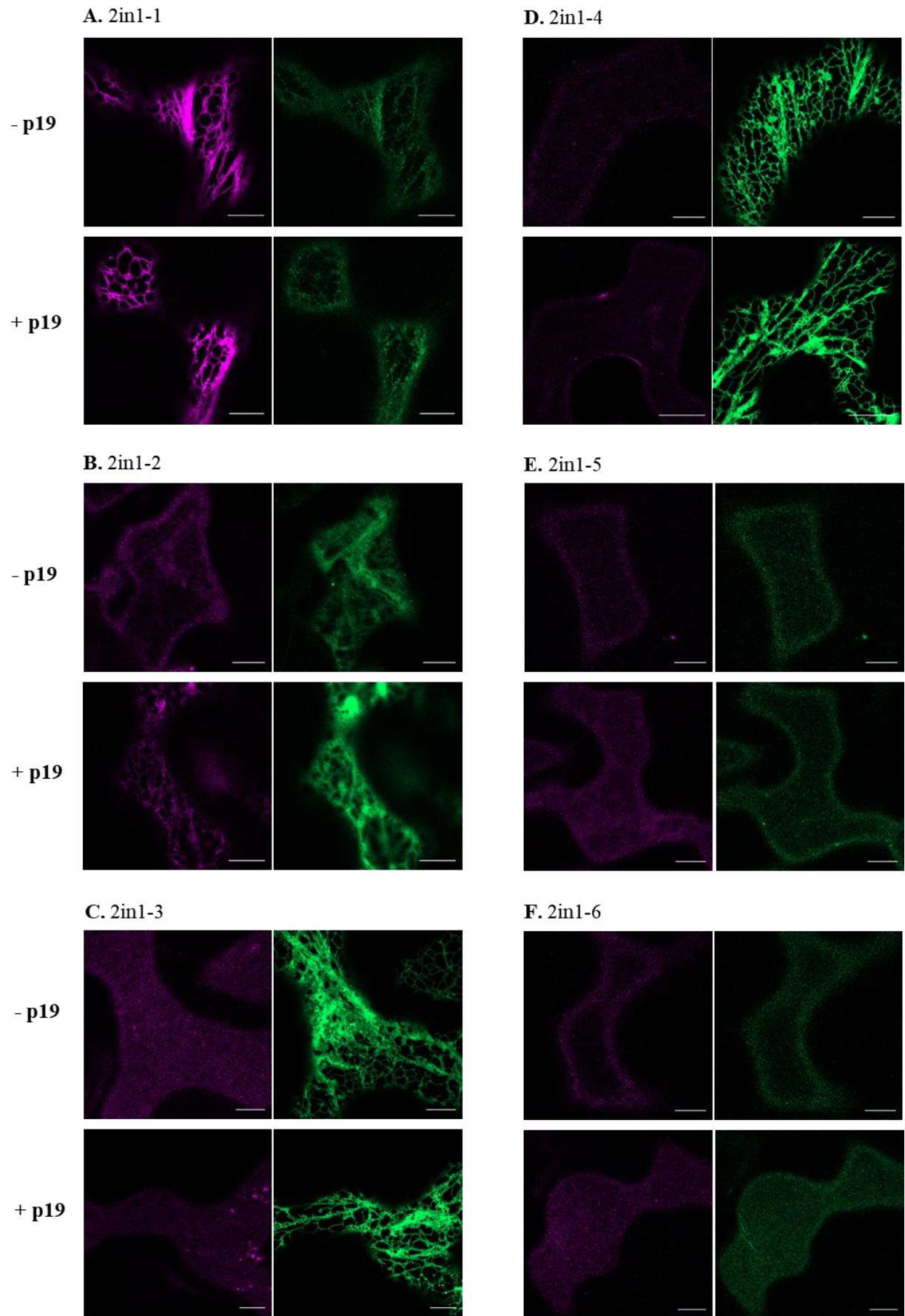


Figure 4.4. Expression of 2in1 vectors with p19.

2in1 vectors were transiently expressed in *Nicotiana benthamiana* with and without p19. Confocal images show the expression patterns in the first and second cassette (red and green respectively). Scale bar = 10 μ m.

4.2.4 Design and transcription of a pMMO polypeptide using self-cleaving peptides

Following the lack of co-expression of pMMO proteins using 2in1 vectors, an alternative gene stacking strategy was investigated using 2A self-cleaving peptides. A quad-cistronic polypeptide was designed containing all pMMO genes separated by different self-cleaving peptides followed by a terminal clover reporter (Figure 4.5 A). This strategy enables the co-expression of all subunits under a single promoter, potentially reducing transgene silencing, as well as improving stoichiometry and providing a more feasible strategy for stable transformation due to the requirement for only a single transformation event. The choice and order of self-cleaving peptides was based on a study that performed a multi-gene expression analysis in mouse lines, which advised against repeating identical self-cleaving peptides within a construct, and reported the highest efficiency for expression of quad-cistronic vectors using 2A peptides in the order T2A, P2A, E2A (Liu *et al.*, 2017). All three 2A sequences comprise a “GSG” linker in their N-terminus which is reported to improve cleavage efficiency (Wang *et al.*, 2015b; Liu *et al.*, 2017). The gene order was maintained as it is present in the bacterial operon (*pmoCAB*), as previous studies have found that the gene order within an operon tends to be optimized for complex assembly in multimeric protein complexes, as proteins adjacent to each other in an operon are more likely to interact (Wells *et al.*, 2016). The *pmoB* version used for the polypeptide was spBtrunc-GFP, containing a plant signal peptide in the place of the bacterial signal peptide to allow targeting to the plant ER. In order to avoid a lower translation efficiency due to codon usage the polypeptide sequence was codon optimized for plants and sent for synthesis. An RFP tag was fused N-terminally using Gateway cloning to be able to see expression and localization of the N-terminal subunit (*pmoC*), yielding the RFP-polypeptide construct. Since translation efficiency decreases towards the C-terminus of long

transcripts due to ribosome drop-off, the presence of clover at the C-terminus was intended to give a measure of translation of the full polypeptide. However, the clover fluorophore was not incorporated into the pMMO complex due to its separation with a self-cleaving peptide, as a means to reduce the amount of protein additions which could potentially interfere with the enzyme's function.

The RFP-polypeptide construct was transiently expressed in tobacco and transcription was assessed via RT-PCR, following RNA extraction at 2 days after infiltration. A band of approximately 3 kb was detected using primers annealing at pmoC and clover, which coincided with the size of the expected transcript, and was absent in non-infiltrated tissue (Figure 4.5 B). CDNA reactions were also carried out in the absence of reverse transcriptase enzyme to rule out the presence of contaminating DNA (noRT controls). PCR was also carried out from the cDNA samples with actin primers, as a control for transcription in both the infiltrated and non-infiltrated tissues.

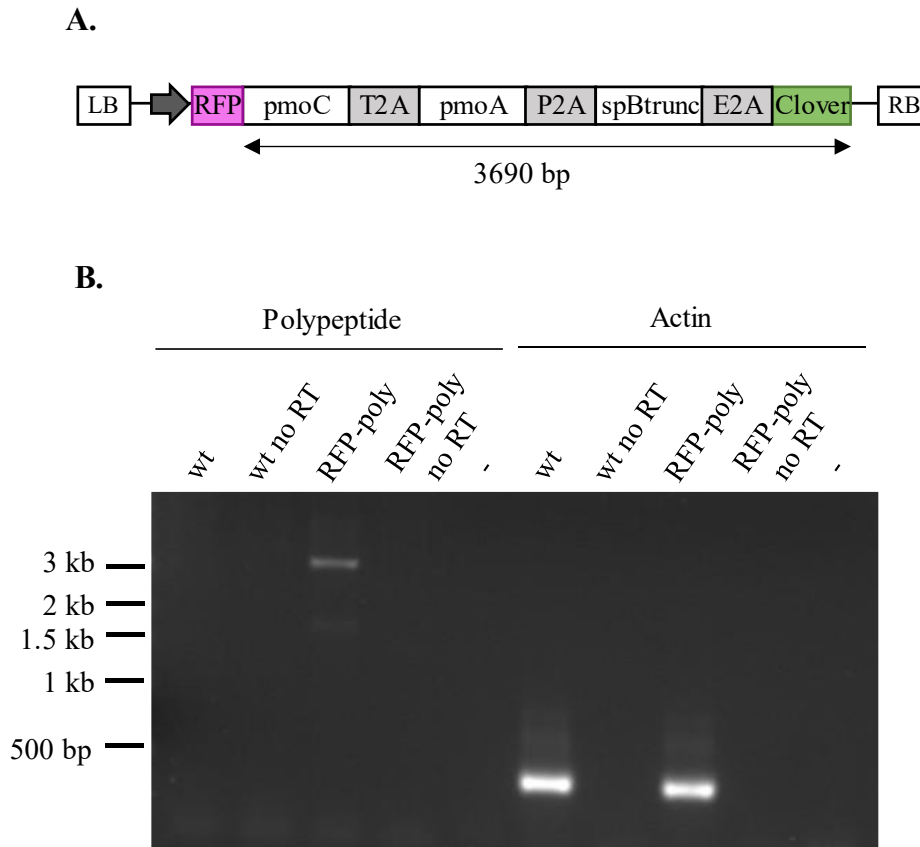


Figure 4.5. Structure and transcription of the pMMO polypeptide.

(A) Schematic diagram of the RFP-polypeptide construct containing three self-cleaving peptides (T2A, P2A and E2A). (B) RT-PCR after transient transformation of tobacco (2 dpi) with RFP-polypeptide. RNA was extracted from a leaf piece in the infiltrated area (RFP-poly) or from a non-infiltrated leaf piece (wt). No RT controls for cDNA reactions without reverse transcriptase are shown. PCR was carried out for the polypeptide transcript (expected size 3,251 bp) and with actin primers (expected size 435 bp).

4.2.5 Co-expression and localization of subunits using 2A self-cleaving peptides

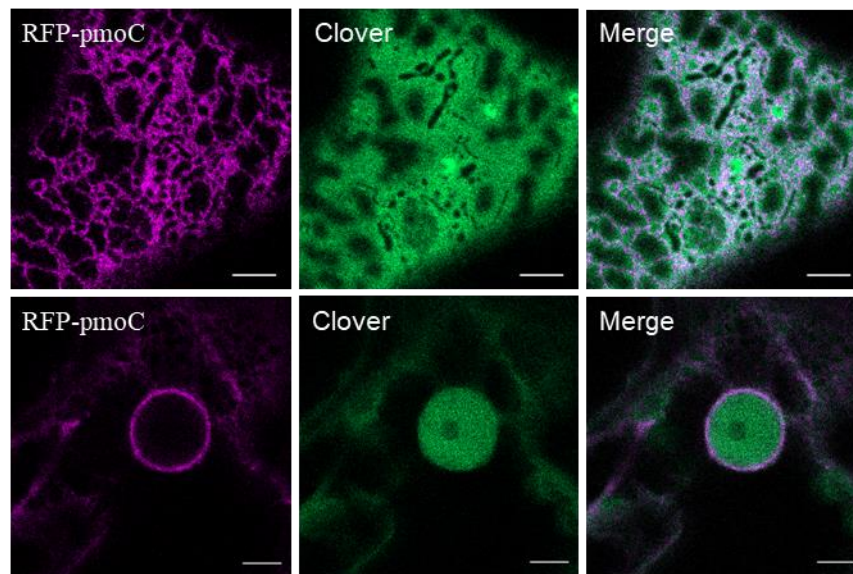
Given that the RFP-polypeptide was transcribed correctly in tobacco, protein translation and subcellular localization were assessed by confocal microscopy. Expression of this construct at 2 days after infiltration showed most cells producing clover in the cytosol, indicating translation of the full length transcript, but very few cells in an ROI contained any red signal from RFP-pmoC. In the cases where an RFP signal was present, this was found in the ER

(Figure 4.6 A). An image of RFP-pmoC labelling the nuclear envelope and clover labelling the nucleoplasm highlights the distinct subcellular localization of both proteins.

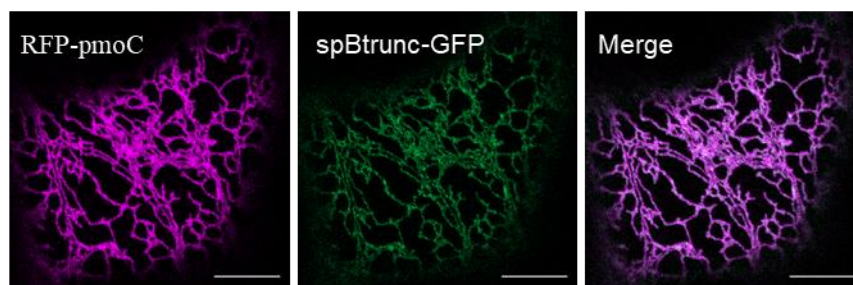
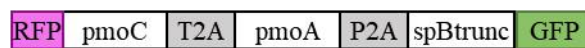
To determine the subcellular localization of pmoA and pmoB within the polypeptide, new constructs were created containing fused fluorophores at both ends: RFP-CAB-GFP (RFP-pmoC-T2A-pmoA-P2A-spBtrunc-GFP) and RFP-CA-GFP (RFP-pmoC-T2A-pmoA-GFP). A summary of all polypeptides created in this chapter is included in appendix 3 (section 9.2), stating the subcellular localization of proteins and their expression levels. Expression levels of RFP-CAB-GFP were similarly low as those of RFP-polypeptide. However, in cells where expression was observed, both RFP-pmoC and spBtrunc-GFP could be seen in the ER (Figure 4.6 B). Moreover, the ER structure did not appear visibly perturbed in the RFP-CAB-GFP and RFP-polypeptide constructs. Expression levels of RFP-CA-GFP were slightly higher, with localization of both RFP-pmoC and pmoA-GFP in the ER. However, most cells expressing this construct presented either perturbations to ER morphology or disruption of the ER network (Figure 4.6 C).

Taken together, the subcellular localisation studies showed that all subunits within the polypeptide were being targeted correctly to the ER and that the terminal clover was being cleaved properly via processing of the E2A self-cleaving peptide, and localized to the cytosol.

A. RFP-polypeptide



B. RFP-CAB-GFP



C. RFP-CA-GFP

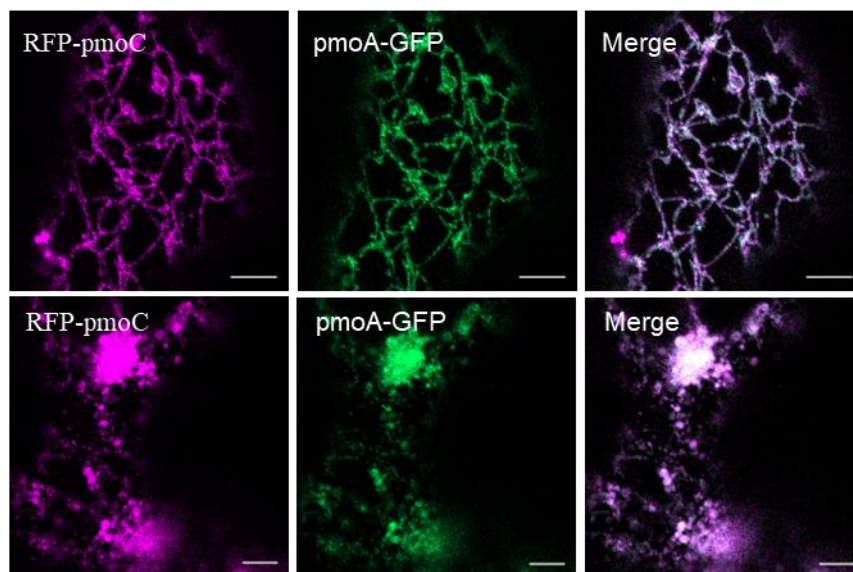


Figure 4.6. Co-expression using 2A self-cleaving peptides.

(A) Transient expression of RFP-polypeptide showing localization of RFP-pmoC in the ER and of clover in the cytosol (bottom: highlight of the differential localization of RFP-pmoC in the nuclear envelope and of clover in the nucleoplasm). (B) Transient expression of RFP-CAB-GFP showing localization of both RFP-pmoC and spBtrunc-GFP in the ER. (C) Transient expression of RFP-CA-GFP showing localization of both RFP-pmoC and pmoA-GFP in the ER, with perturbation to ER morphology . Scale bar = 5 μ m.

4.2.6 Quantifying expression levels of pMMO polypeptides

The co-expression strategy using self-cleaving peptides appeared successful both in terms of simultaneously expressing two or more subunits, and localizing all subunits correctly at the ER. However, expression levels are also an important factor in transgene production as low amounts of the transgenes may not be sufficient to produce the desired physiological result. As mentioned previously, the expression levels of RFP-polypeptide in tobacco appeared to be low. A quantification method was developed using low magnification images and defined imaging conditions, to compare the expression levels of polypeptides with that of single subunits (see materials and methods). The signal intensity of RFP-tagged constructs was extracted from 3 different plants (biological repeats) and 20 ROIs each (technical repeats). Expression levels of RFP-polypeptide were compared between RFP-CAB (which does not contain clover), RFP-CA, pmoC-RFP, pmoA-RFP, spBtrunc-RFP_{hdel} and RFP-HDEL (an ER marker used as a reference for high expression levels). The imaging conditions were set up with a low laser power and gain to reduce the amount of saturated pixels in the highly expressing constructs, and therefore increase the dynamic range for the detection of differences. Under these conditions, ER signal in RFP-polypeptide and RFP-CAB, which were the lowest expressing constructs, was almost undetectable (Figure 4.7). The RFP intensity levels extracted from the ROIs was indistinguishable to those from a non-infiltrated leaf section (labelled as autofluorescence). In contrast, the expression levels of RFP-CA and the single subunits were comparable to that of RFP-HDEL. RFP-CA and pmoC-RFP induced disruption of the ER network in almost 100% of expressing cells, whereas pmoA-RFP induced perturbations to the

ER structure only in some cells, whilst spBtrunc-RFPhdel did not produce any visible perturbations to the cells.

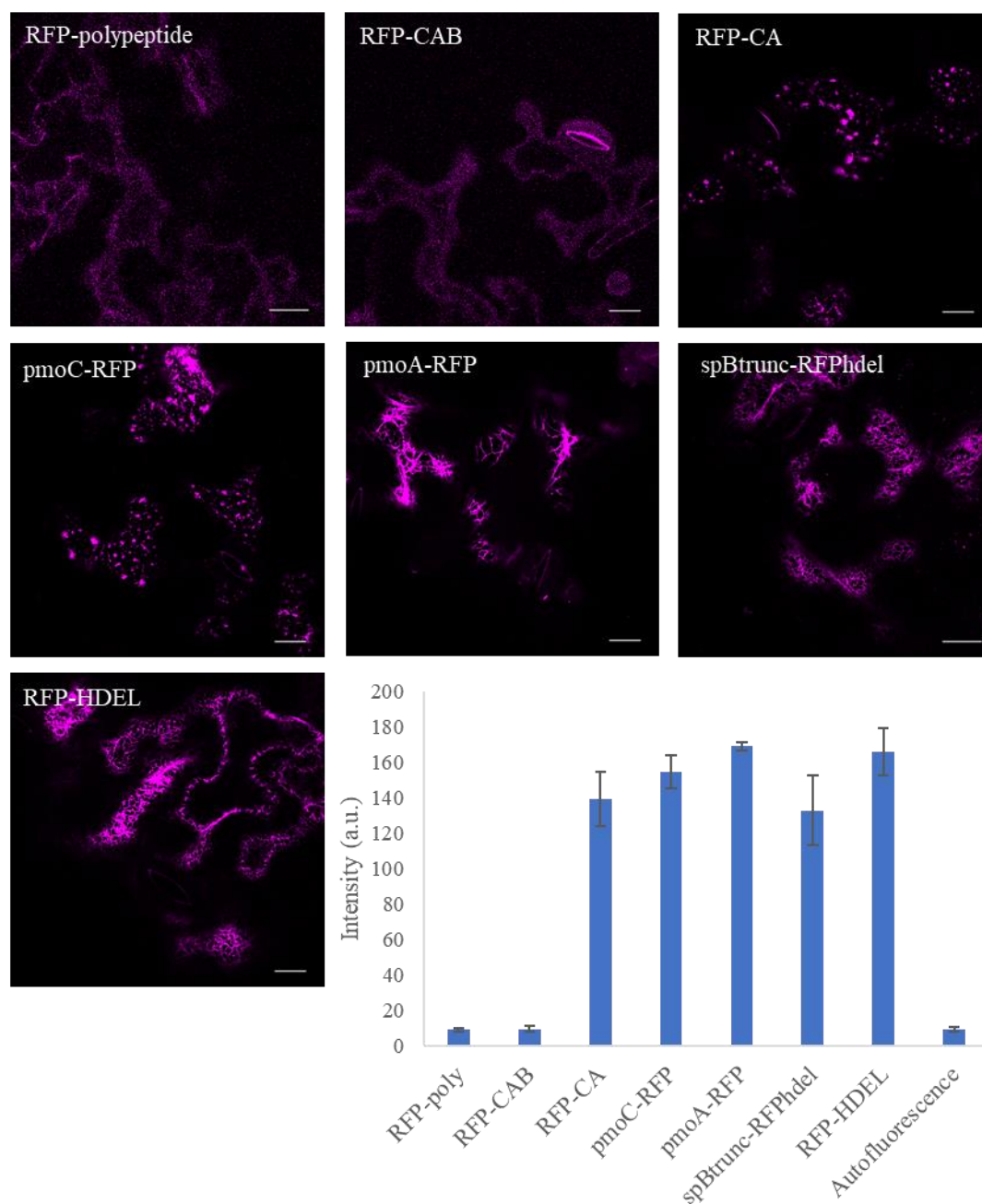


Figure 4.7. Expression levels of RFP-tagged polypeptides and single subunits.

Example of low magnification images used for the quantification are shown for each construct. All constructs were infiltrated at OD₆₀₀ 0.1 and imaged under constant conditions in three separate plants (biological repeats) and 20 ROIs per plant (technical repeats). A threshold was applied to extract the average signal intensity excluding the background (see materials and methods) The graph shows the mean and standard error of three biological replicates for each

construct and for a non-infiltrated leaf section (autofluorescence). Laser power = 0.6%, gain = 650 V. Scale bar = 100 μ m.

The suppressor of gene silencing p19 was co-expressed as a strategy to boost expression levels. In this case, co-expression with p19 resulted in an increase of RFP-polypeptide levels, indicating that PTGS is fully or partly responsible for the polypeptide's low expression in plants. The average signal intensity levels were measured as described previously, using 5 biological replicates and 20 ROIs per replicate (Figure 4.8). On average the RFP signal from RFP-polypeptide without p19 was 34 ± 5 (mean $\pm$ s.e.m, n=5) and with p19 was 107 ± 14 (mean $\pm$ s.e.m, n=5) measured on an 8 bit depth (0-255). The GFP signal, from the terminal clover in the polypeptide, was also considerably increased with p19, changing from an average of 41 ± 3 (mean $\pm$ s.e.m, n=5) without p19 to 94 ± 8 (mean $\pm$ s.e.m, n=5) with p19.

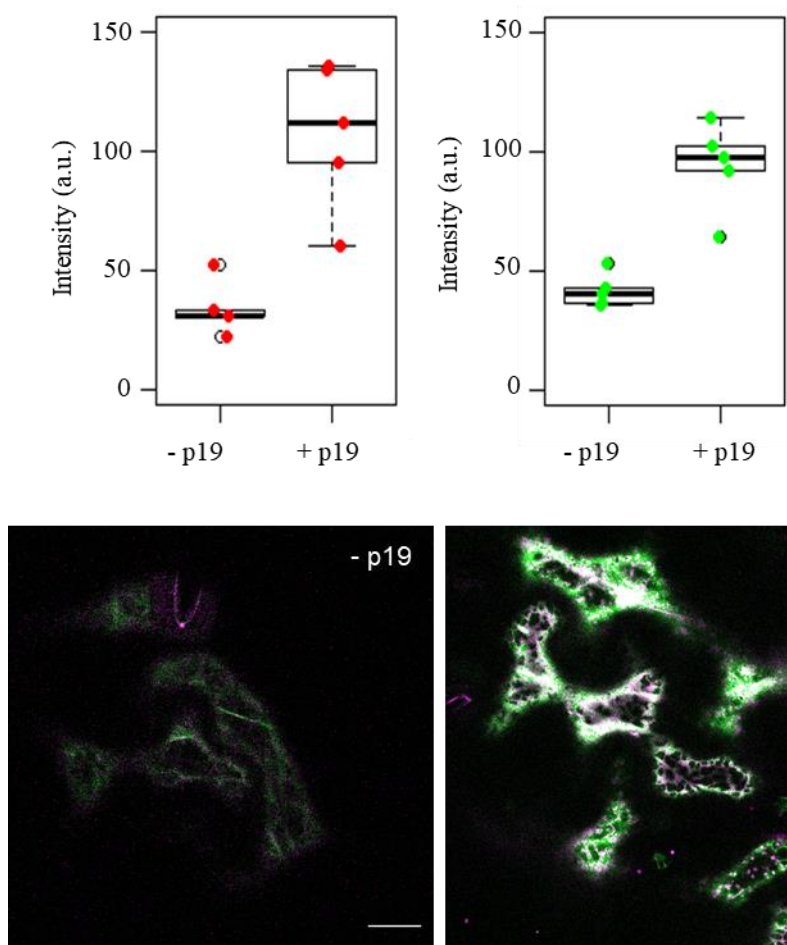


Figure 4.8. Expression levels of RFP-polypeptide with and without p19 in *Nicotiana benthamiana*.

Average fluorescence intensity following transient expression of RFP-polypeptide from 5 plants (20 ROIs each), in the red and green channels. Example images of RFP-polypeptide with and without co-expression with p19 are shown. Laser power = 0.7%, gain = 700 V (green channel) and 750 V (red channel). Scale bar = 100 μ m.

4.2.7 Protein-protein interactions in pMMO polypeptides

To test whether the subunits might assemble on the ER membrane in a similar manner to when co-expressed from separate plasmids, FRET-FLIM was carried out to test protein-protein interactions with the polypeptide constructs. Interactions were tested between RFP-pmoC and pmoA-GFP in the RFP-CA-GFP construct, comparing the GFP lifetime value to that of pmoA-GFP alone. The increase in expression levels achieved by co-expression with p19 permitted FRET-FLIM on the RFP-CAB-GFP construct, and the GFP lifetime was compared to that of spBtrunc-GFP_{hdel} alone. Figure 4.9 shows representative images of the nuclear envelope in

cells expressing RFP-CA-GFP and RFP-CAB-GFP which were used for FRET-FLIM, along with the analysis through the software. The fluorescence lifetime decay curves were fitted to single exponential functions from which a lifetime value was extracted. Curves which presented a good fit (χ^2 between 0.9 and 1.2) were used for analysis, and measurements were taken from three or four different plants for each combination. The average lifetime of pmoA-GFP alone (2.34 ± 0.01 ns) was significantly reduced when present in RFP-CA-GFP (2.07 ± 0.03 ns). The reduction in lifetime of 0.27 ns is indicative of an interaction between RFP-pmoC and pmoA-GFP in the polypeptide (Figure 4.10). On the other hand, the average lifetime of spBtrunc-GFP_{hdel} alone (2.45 ± 0.01 ns) did not differ from the GFP lifetime in RFP-CAB-GFP (2.45 ± 0.01 ns), which is indicative of no interaction between the fluorophores of RFP-pmoC and spBtrunc-GFP (Figure 4.10). Both results agree with the interactions and proposed topology shown in section 3.2.4 with the single subunits, where pmoB termini are facing the lumen of the ER and therefore don't show an interaction via FRET-FLIM with pmoC and pmoA, whose termini face the cytosol and interact with each other via FRET-FLIM.

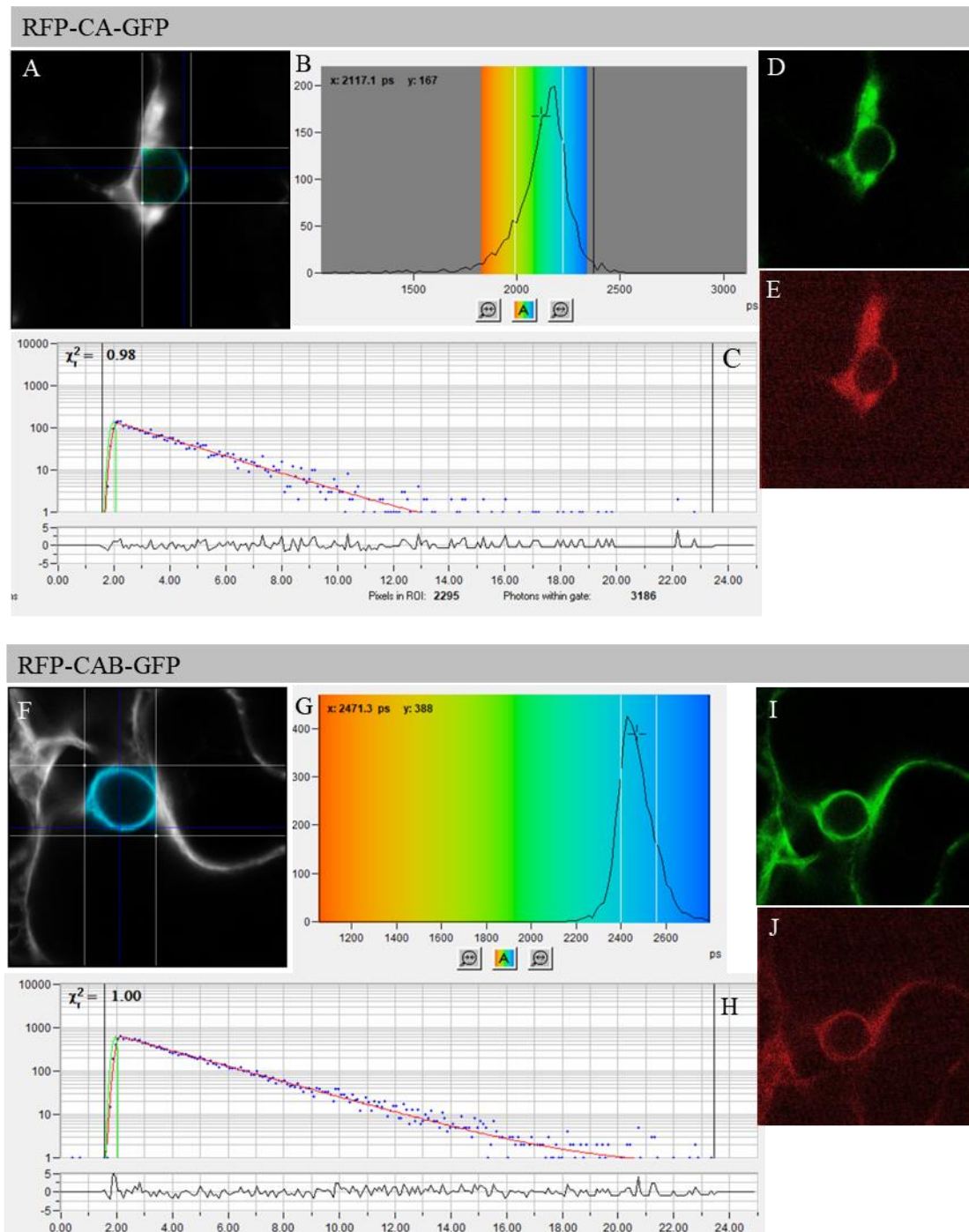


Figure 4.9. FRET-FLIM analysis of polypeptides.

(A, F) Pseudo-colored lifetime maps for nuclear envelopes expressing RFP-CA-GFP and RFP-CAB-GFP, respectively, showing the region selected for the analysis. (B, G) Distribution of the lifetime values for pixels within the selected region, color-coded according to the lifetime values with blue colors indicating longer lifetimes. (C, H) Fluorescence lifetime decay curve of a single pixel within the region, showing the χ^2 value. (D, E, I, J) Raw FRET-FLIM images.

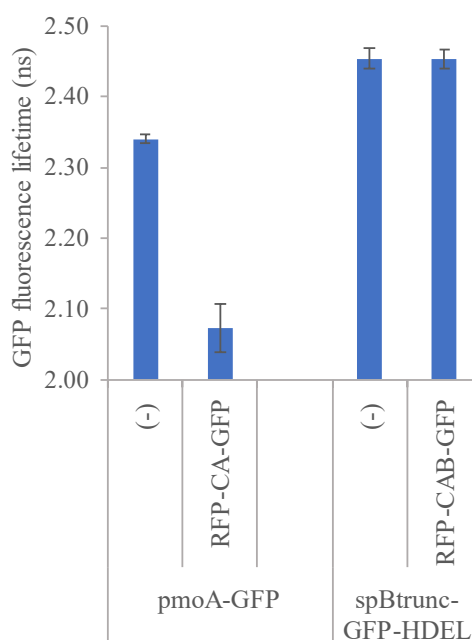


Figure 4.10. Fluorescence lifetime values for pmoA-GFP and spBtrunc-GFP_{HDEL} alone or within polypeptides.

Average lifetime values with s.e.m. for pmoA-GFP, RFP-CA-GFP, spBtrunc-GFP_{HDEL} and RFP-CAB-GFP with n (number of cells)=(8, 6, 6, 6), respectively, taken from 3 or 4 different plants.

4.2.8 Ratiometric expression of pMMO subunits in polypeptides

Obtaining a specific and robust fixed stoichiometry of proteins is advantageous for many applications using recombinant systems, particularly for protein complex formation. Using standard confocal microscopy it is not straight forward to determine the absolute protein quantities as fluorescence levels are typically expressed in arbitrary units. However, by keeping the same imaging conditions, it is possible to determine how constant the relative amounts are in a sample, i.e. if the proteins are expressed in ratiometric quantities. The ratio of fluorescence intensities was studied in the RFP-CAB-GFP construct at OD₆₀₀ 0.1, co-expressed with p19 to obtain higher expression levels in *Nicotiana benthamiana*. Low magnification images (15 ROIs) containing several cells each were acquired, and the intensity values in the green and red channels were measured using the ImageJ freehand selection tool, on a cell-by-cell basis. The green/red ratio of intensity values was compared to that obtained by co-expressing two

subunits (pmoC-RFP and GFP-pmoA) via co-infiltration of separate plasmids at OD₆₀₀ 0.1 each.

Example images of the merged channels show a more consistent ratio of intensities with the RFP-CAB-GFP construct compared to the co-infiltration strategy (Figure 4.11 A and B). Quantification of the green/red ratio in over 30 cells reflects this by showing a smaller variation in the ratio in RFP-CAB-GFP, where the $\log_2(\text{green/red})$ of RFP-CAB-GFP has a mean of -0.07 and standard deviation of 0.36 (with 36 cells), and pmoC-RFP + GFP-pmoA co-infiltration has a mean of -0.8 and standard deviation of 1.2 (with 32 cells), as well as a smaller range and interquartile range in RFP-CAB-GFP (Figure 4.11 C).

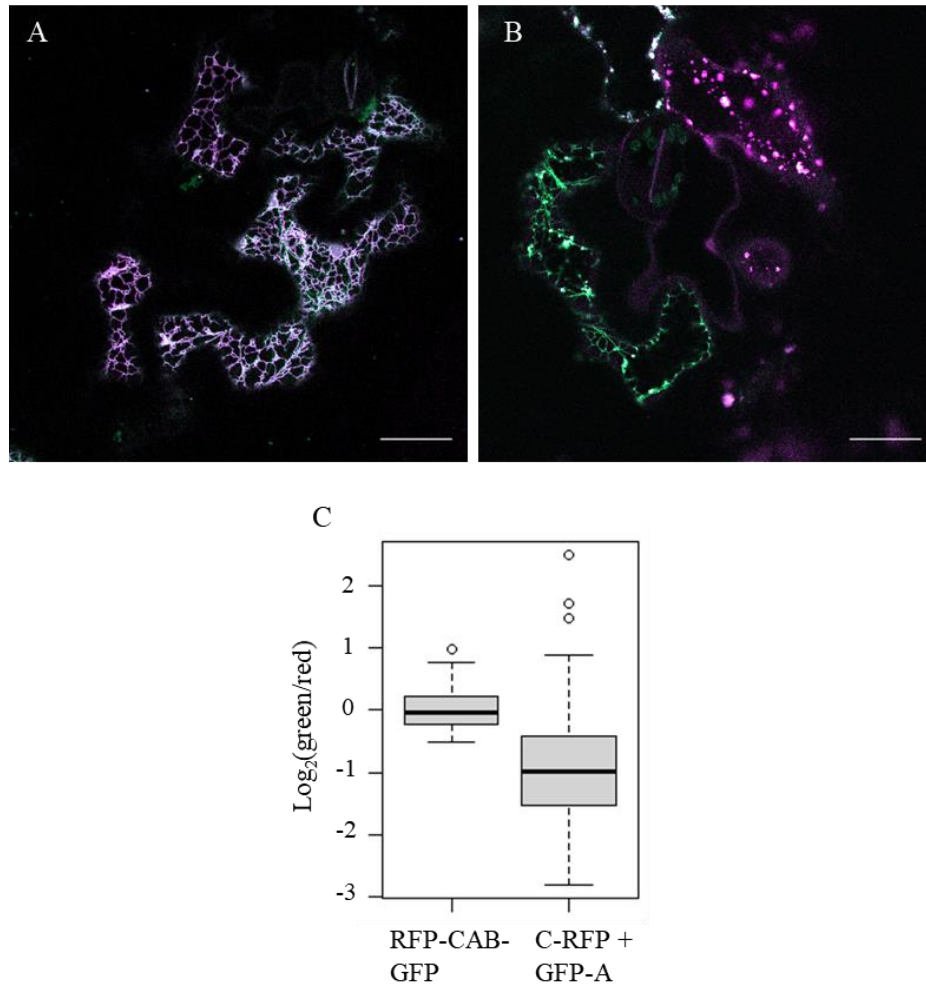


Figure 4.11. Variation in the ratio of green/red fluorescence intensity levels.

(A) Merged channel image of transiently expressed RFP-CAB-GFP (OD₆₀₀ 0.1) + p19 in *Nicotiana benthamiana*. (B) Merged channel image of transiently expressed pmoC-RFP (OD₆₀₀ 0.1) + GFP-pmoA (OD₆₀₀ 0.1) in *Nicotiana benthamiana*. (C) Log₂ of the green/red ratio of fluorescence intensity (using 36 cells for RFP-CAB-GFP, and 32 cells for pmoC-RFP + GFP-pmoA). Scale bar = 100 μm .

4.3 Discussion

4.3.1 Evaluation of co-expression using 2in1 vectors

Plant 2in1 vectors (Hecker *et al.*, 2015) were tested in this study with the aim of co-expressing two subunits from the same T-DNA. However, 2in1 vectors containing the pMMO subunits did not yield dual expression, instead only one cassette expressed at a time, even when varying the position of the tag and the combination of subunits (Figure 4.2). Interestingly, the N-terminally tagged pmoC expressed in the first cassette in 2in1-1, and C-terminally tagged pmoA expressed strongly in the second cassette in 2in1-3 and -4. However, when combining these two cassettes creating 2in1 5, expression seemed to be lost in both cassettes, indicating some sort of negative interaction in the presence of both subunits. Single cassette 2in1s where the strongly expressing pmoA was not included in the second cassette, yielded expression of the first cassette, confirming the hypothesis of a silencing or inhibitory mechanism occurring in the presence of more than one subunit (Figure 4.3).

One type of silencing which is often induced with the expression of foreign proteins is post-transcriptional gene silencing (PTGS), which targets the transgene mRNA for degradation via the formation of small interfering RNAs (siRNAs) (Vaucheret *et al.*, 2001). Co-expression with p19 is effective at blocking this silencing pathway by sequestering siRNAs, and is often used to boost transgene expression levels (Voinnet *et al.*, 2003). However, co-expression of 2in1s with p19 did not lead to a change in expression patterns, therefore a different type of silencing mechanism must be taking place. Another type of silencing is transcriptional gene silencing (TGS), which occurs via the methylation of regulatory regions or coding sequences, between others, inhibiting transcription (Vaucheret and Fagard, 2001). Many mechanisms could trigger TGS and further studies would be required to explore the silencing in 2in1 constructs containing pMMO proteins. A potential trigger in this case could be the interaction between the pMMO coding sequences and the tandem 35S promoter structure in the vectors.

Regardless of the potential mechanisms occurring, 2in1 vectors did not prove to be useful for co-expression of pMMO proteins. In *Nicotiana benthamiana* some co-expression could be seen with 2in1-1, however, expression of the second cassette was very faint, and this result could not be reproduced in tobacco. Therefore, alternative co-expression methods were explored.

4.3.2 Co-expression and ER targeting with 2A self-cleaving peptides

2A self-cleaving peptides are an increasingly popular method for gene stacking in eukaryotic cells, with the use of a single promoter for co-expression providing many benefits. Expression of the first element of the polypeptide (RFP-pmoC) was correctly located to the ER, and expression of the final element of the polypeptide (clover) in the cytosol (Figure 4.6). RT-PCR showed transcription of the full length mRNA for the RFP-polypeptide, containing all pMMO subunits and clover separated by self-cleaving peptides (Figure 4.5). Given the mechanism of polypeptide synthesis from a single mRNA molecule, and continuous translation with proteins being separated via a ribosomal skip, this result suggested that translation was occurring throughout the polypeptide, in which case all three subunits would be synthesized. To test whether the polypeptide system would produce correct sequential ER targeting of pmoA and spBtrunc, double fluorophore fusion vectors RFP-CAB-GFP and RFP-CA-GFP were constructed, and showed that spBtrunc-GFP and pmoA-GFP were expressed and localized correctly to the ER (Figure 4.6).

Previous studies have shown that the use of 2A peptides can be combined with targeting proteins to different subcellular compartments in plant cells, including the ER (El Amrani *et al.*, 2004; Samalova *et al.*, 2006; Burén *et al.*, 2012). Samalova and co-workers proposed a model in which the 2A peptide was necessary when targeting two proteins sequentially to the ER lumen: in the absence of the 2A sequence the first protein would be correctly targeted to the ER via its signal anchor or signal peptide, whereas the signal peptide of the second protein would be translocated in the incorrect direction with its amino terminus facing the lumen, and therefore act as a stop-transfer sequence which was not cleaved by signal peptidase, leading to

a misfolded fusion protein subject to degradation (Figure 4.12 A); on the other hand in the presence of a 2A sequence between the proteins, structural characteristics of the 2A sequence would allow the second signal sequence to flip orientation, and enter the translocon with its C-terminus facing the lumen, allowing the translocation of the second protein into the lumen (Figure 4.12 B) (Samalova *et al.*, 2006).

In the case of the pMMO polypeptide, although the first three proteins localize to the ER, pmoC and pmoA are membrane proteins with their C-termini facing the cytosol. Therefore, the translocation of these proteins targeted sequentially to the ER would differ from the model in Figure 4.12. Potentially, the first 2A peptide following pmoC would be translated in the cytosol without translocating through Sec61. Following cleavage, the first transmembrane of pmoA (acting as a signal anchor) would orient itself correctly and translocate through the ER membrane. The same process would occur between pmoA (with its C-terminus in the cytosol) and spBtrunc, and finally, the spBtrunc C-terminus is located in the ER lumen and cleavage of the last self-cleaving peptide would result in clover localizing to the cytosol. Confocal images of RFP-polypeptide show clover being localized exclusively to the cytosol, indicating that E2A is cleaving properly. However, western blots could be carried out to evaluate the cleavage efficiency of the first two self-cleaving peptides by detecting the presence of cleaved moieties and uncleaved products with a larger size.

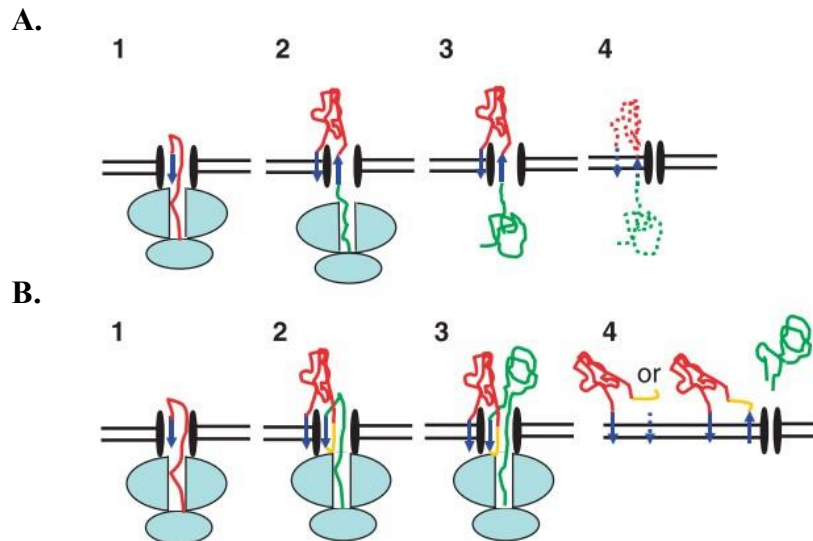


Figure 4.12. Role of 2A sequence in sequential ER targeting.

Model proposed by (Samalova *et al.*, 2006) where (A) no self-cleaving peptide between two proteins targeted to the ER lumen leads to the second signal peptide (blue arrow) acting as a stop-transfer signal, preventing translocation, and leading to an unstable fusion protein, and (B) the presence of a self-cleaving peptide (yellow line) allows the second signal peptide to approach the translocon with the correct orientation (N-terminus facing the cytosol), and translocation of the second protein (green) into the ER. Figure taken and cropped from (Samalova *et al.*, 2006).

4.3.3 Interactions and stoichiometry in 2A polypeptides

Protein-protein interactions in polypeptides were tested via FRET-FLIM and reflected the interactions shown between the single subunits in Chapter 3. A significant decrease in the GFP lifetime in RFP-CA-GFP compared with pmoA-GFP was indicative of an interaction between RFP-pmoC and pmoA-GFP, and is in accordance with the topology where both termini are facing the cytosol. Conversely, no interaction was observed between RFP-pmoC and spBtrunc-GFP in FRET-FLIM analysis of RFP-CAB-GFP. This was also consistent with topology analysis where both pmoC termini are predicted to be in the cytosol, and spBtrunc termini in the ER lumen. Thus no detectable interaction between the subunits is expected via a FRET-derived method as the fluorophores are positioned on opposite sides of the ER membrane. These results suggest that the topology and interactions of the subunits in polypeptide constructs do not differ from when the subunits are expressed from individual plasmids.

By calculating the green/red ratio in fluorescence intensity, it was found that RFP-CAB-GFP expresses at least two of the pMMO subunits in a ratiometric manner, with the green/ratio having relatively low cell-to-cell variation, in comparison with co-infiltrating two subunits from separate plasmids (Figure 4.11 C). A1:1:1 stoichiometry of the subunits would be an ideal scenario to achieve the maximum amount of assembled pMMO. The actual stoichiometry between subunits cannot be calculated with standard confocal microscopy, as fluorescence intensity levels are correlated not only with protein concentration but also with the fluorescent protein spectral properties, quantum yield, and the imaging setup (Cherkas *et al.*, 2018). Other methods based on fluorescent imaging, such as fluorescence correlation microscopy, are able to provide information regarding absolute protein numbers (Verdaasdonk *et al.*, 2014). However, it can be seen that by co-expressing with p19, substantial levels of expression of both subunits are present.

Besides from attaining higher expression levels of RFP-poly and RFP-CAB-GFP with p19, it could be seen that the ER was present as a contiguous network, not showing apparent damage to the cell, as was observed previously with pmoC alone and with RFP-CA. The results in this chapter therefore suggest that the use of 2A self-cleaving peptides, combined with the PTGS suppressor p19, is a viable strategy to co-express the pMMO subunits in the plant ER at high level, consistent stoichiometry, and potentially lower disruption to ER structure.

5 Impact of pMMO on ER structure and dynamics

5.1 Introduction

5.1.1 Structure and movement of the plant ER

The ER is a vast multifunctional organelle in the cell which presents a complex and dynamic network-like structure. The network is contiguous with the nuclear envelope and extends throughout the cytoplasm up to the plasma membrane, presenting multiple contact sites with cellular organelles (Friedman and Voeltz, 2011; Sparkes, 2018). The ER is structured into two main elements, tubules and cisternae: tubules are thin highly curved structures which connect with each other forming “three-way junctions” and with cisternae, whereas cisternae (also referred to as sheets) are flat regions with a larger surface area (Kriechbaumer and Brandizzi, 2020) (Figure 5.1). Recent studies carried out in mammalian cell lines using superresolution microscopy, have contested the classical view on cisternae, suggesting that sheets are composed of densely packed matrices of tubules (Nixon-Abell *et al.*, 2016), or that they contain dynamic nanosized holes (Schroeder *et al.*, 2019).

The plant ER architecture is thought to be formed and maintained by the action of three classes of ER-shaping proteins (Kriechbaumer and Brandizzi, 2020). The first class is the reticulon family: reticulon proteins are integral membrane proteins that participate in the formation and maintenance of tubules by inducing membrane curvature through their W-shaped topology and through oligomerization (Sparkes *et al.*, 2010). Reticulon proteins are also found to localize to ER tubules and induce a tubular network when over-expressed in yeast or mammalian cells (Voeltz *et al.*, 2006). The second class corresponds to the Lunapark family: overexpression of *Arabidopsis* Lunapark 1 and Lunapark 2 leads to an increase in cisternal area in tobacco, and *Arabidopsis lnp1* and *lnp2* knockdown mutants show a decrease in the cisternal area, suggesting a role of the Lunapark family in cisternal formation in plants (Kriechbaumer *et al.*,

2018). Nevertheless, alternative localization patterns and mechanistic actions of the Lunapark proteins have been proposed in yeast and mammalian cells (Chen *et al.*, 2012, 2015), and recently also in plants (Sun *et al.*, 2020). Thirdly, atlastin proteins in mammals and Sey1p in yeasts are dynamin-like GTPases which act as ER fusogens, mediating homotypic-membrane fusions leading to the interconnection of tubules (Brandizzi, 2021). RHD3 (root hair deficient mutant) is a plant homologue of mammalian atlastins and is thought to play a similar role in the fusion of tubules (Brandizzi, 2021).

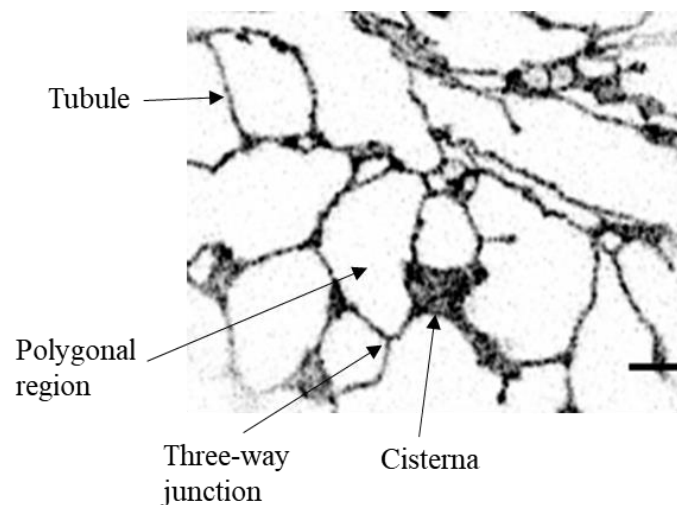


Figure 5.1. Structural elements of the ER network.

Subsection of a cell expressing an ER-localized GFP, showing the different structural elements which make up the ER architecture. Tubules are thin, elongated structures which connect with each other at three-way junctions or with cisternae, whereas cisternae are flat regions with a larger surface area. The tubule and cisternal edges connect with each other enclosing polygonal regions. Scale bar = 2 μm .

The ER structure is highly dynamic, and as a result, tubules and cisternae undergo rapid remodeling and interconversion. Using live-cell confocal imaging, changes such as the extension and retraction of tubules, reshaping and merging of cisternae, formation of new interconnections, polygonal ring formation and ring closure, amongst others, can be visualized. The ER overlays the actin cytoskeleton and its movement in plants has been shown to be based on actin and myosins XI-K and XI-2 (Sparkes *et al.*, 2009; Ueda *et al.*, 2010). Image analysis software developed to quantify ER movement revealed the presence of more static or

“persistent” nodes, which do not undergo dynamic changes in the time-lapse images studied, as well as general remodelling of the network around these persistent nodes, and a faster type of movement occurring over actin filament bundles termed cytoplasmic streaming or bulk flow (Sparkes *et al.*, 2009; Ueda *et al.*, 2010; Pain *et al.*, 2019). The types of ER movement and quantification methods available have been recently reviewed (Pain and Kriechbaumer, 2020). The dynamic nature of the ER is the basis on which the organelle adapts to alterations in the cellular environment. For example, the ER structure in plants has been shown to be highly sensitive to changes such as heat-shock, drug treatments, and transient protein over-expression, amongst others (Saint-Jore *et al.*, 2002; Runions *et al.*, 2006; Pain *et al.*, 2019). Significant changes were observed in the proportion of cisternae to tubules, as well as changes in dynamics, under those conditions. Moreover, stable differences in the proportion of tubular and cisternal ER are observed in different cell types and development stages (Kriechbaumer and Brandizzi, 2020), suggesting a possible function for the different structural conformations of the ER. A possible link between structure and function of the ER has been suggested in mammalian cells, where cells with high secretory demands, such as pancreatic cells, are seen to contain a high proportion of cisternae, whereas muscle cells, which require the rapid transport of calcium for contraction, are enriched in tubules (Friedman and Voeltz, 2011). Although a concrete connection between ER structure and function has not been established, it seems likely that maintaining ER homeostasis is important to retain the multiple cellular processes in which the organelle plays a central part.

5.1.2 AnalyzER: a tool for the quantitative analysis of ER structure and dynamics

A vast literature exists addressing the nature and changes to ER structure and increasingly the study of dynamics. Some studies have developed image analysis tools which quantify different aspects of the ER network, such as the analysis of speed and persistent points (Sparkes *et al.*, 2009; Ueda *et al.*, 2010), modelling of the ER dynamics and network geometry in the context

of underlying biophysical forces (Lin *et al.*, 2014, 2017), or modelling ER architecture through the quantification of ER junctions (Spits *et al.*, 2021), amongst others. However, most studies still report changes which are based solely on visual inspection of the sample and are represented with example images. Due to the relatively large natural cell-to-cell variation already present in ER structure (Figure 5.2), visual inspection may introduce subjectivity and error in the analysis. Moreover, some less evident changes in structure may go undetected. Therefore, automated processing combined with statistical analysis for a large number of cells provides a more accurate description of quantitative changes to the ER.

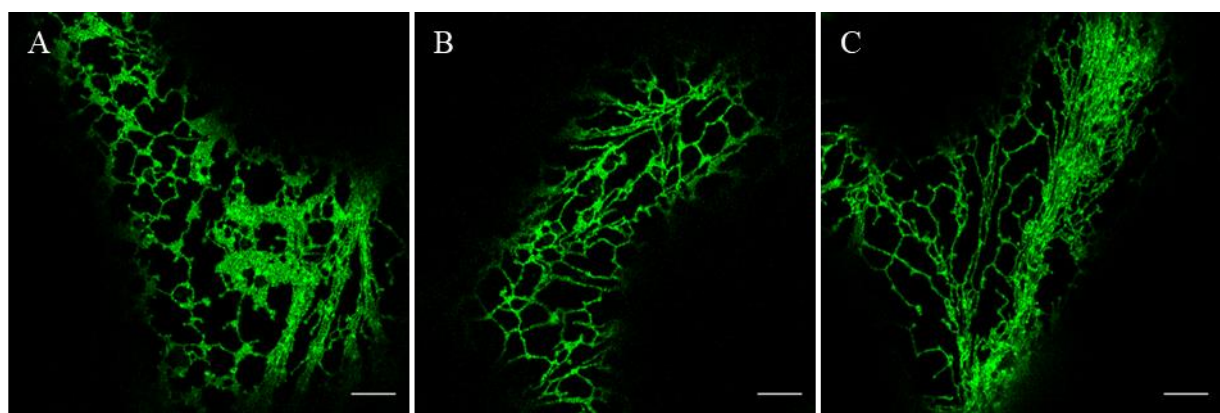


Figure 5.2. Natural cell-to-cell variation in ER morphology.

Airyscan images of tobacco cells expressing GFP-HDEL from the same ROI showing different ER morphologies: (A) a high abundance of cisternae, (B) an intermediate proportion of tubules and cisternae, (C) a highly tubular network with a stream of appressed tubules. Scale bar = 5 μm .

The AnalyzER software (Pain *et al.*, 2019) was designed to quantify changes in the ER using live cell imaging, performing an automated analysis of time series and using a comprehensive set of metrics in order to describe the complex ER structure and dynamics. The software pre-processes the images through background subtraction and filtering, and then segments the ER network using an intensity-based threshold combined with intensity-independent phase congruency enhancement to include dimmer structures such as thin tubules. The network is then converted into a single pixel-wide “skeleton”, where tubules are represented as lines, and cisternae are defined based on an area cut-off. During subsequent processing, an array of

information is extracted which is detailed in the software manual, including metrics on tubule edges (such as length and width of tubules), tubule morphology, cisternae morphology, cisternae texture, polygonal regions, speed, and persistency (Pain *et al.*, 2019). A statistics package is also implemented to facilitate the analysis of the data, including multivariate analysis to study global changes combining several metrics.

AnalyzER has been used to study the phenotypes induced by plant ER-shaping proteins such as *Arabidopsis* reticulon1 and *Arabidopsis* Lunapark 1, and the effects of drug treatments and heat-shock (Pain *et al.*, 2019). It has also been used to characterize the network structure at ER-plasma membrane contact sites (EPCSs) in *Nicotiana benthamiana* (Zang *et al.*, 2021) and to study the impact of glycosylation on ER structure in mammalian cells (Kerselidou *et al.*, 2020).

The aim of this chapter was to study the impact of recombinantly expressing the pMMO polypeptide - producing all three subunits in the ER membrane - on the structure and dynamics of the organelle in tobacco cells. Due to the involvement of the ER in numerous cell functions, ranging from protein production and transport, to lipid biosynthesis and calcium storage, amongst others, and its extensive contact sites with other organelles, maintaining the cell's natural ER architecture is critical to avoid perturbation of cellular processes and function. This is important for a bioengineering strategy where the complex needs to function alongside normal cell physiology. It is therefore important to determine the extent (or not) of ER disruption caused by transgene expression. The ER architecture of cells expressing pMMO was compared to “wild type” cells expressing luminal ER markers, and to cells expressing the subunits individually. Comparisons were carried out complementing visual inspection with an automated quantitative analysis using AnalyzER.

5.2 Results

5.2.1 Evaluation of the impact of polypeptides on the structure of the ER

In chapter 4 it was shown that the ER-targeted pMMO subunits could be simultaneously co-expressed in tobacco using 2A self-cleaving peptides. The expression levels of polypeptides containing all three subunits were low, but were successfully boosted via co-expression with p19. In addition to obtaining sufficient expression levels, the impact on ER integrity is an important consideration for the recombinant system, as disruption of the ER network might adversely affect function and potentially lead to tissue necrosis, limiting methane detoxification *in vivo*.

The levels of ER perturbation are dependent both on intrinsic properties of the proteins and on the protein dosage, where very high recombinant expression levels often lead to ER disruption. Therefore to conduct a valid evaluation of ER perturbation induced by polypeptides and to compare with that induced by the single subunits, images with similar expression levels were used. In order to obtain cells with similar expression levels, the polypeptide constructs (RFP-polypeptide, RFP-CAB and RFP-CAB-GFP) were co-expressed with p19 in *Nicotiana benthamiana*. However, at 48 hs after infiltration the single subunits infiltrated at OD₆₀₀ 0.1 had a very high signal and produced image saturation, therefore conditions of low laser power and gain (laser power 0.6% and gain 600 V) were used to avoid saturation. Between 14 and 20 ROIs with an area of 133 μm^2 and encompassing multiple cells each were taken for each construct to count a significant amount of cells (Figure 5.3). The RFP intensity levels were quantified from a thresholded area as described in materials and methods, and the average intensity levels for the selected ROIs were found to be similar between constructs (Figure 5.4 A). Individual cells were counted and classified as having either a contiguous or a disrupted ER. At one end, spBtrunc-RFP showed no obvious ER disruption, whereas at the other end, pmoC-RFP showed severe disruption in all cells (Figure 5.4 E and G). The polypeptides and pmoA-RFP presented intermediate percentages of contiguous and disrupted ER in different

cells (Figure 5.4 B-D, F). This indicated that tobacco cells are able to harbour polypeptide constructs with high levels of expression, but with varying number of cells showing no disruption to the ER morphology. Interestingly, polypeptides containing all three pMMO subunits were not as highly damaging to the cell, as opposed to expressing pmoC or pmoA subunits individually. Indeed pmoC caused a necrotic phenotype.

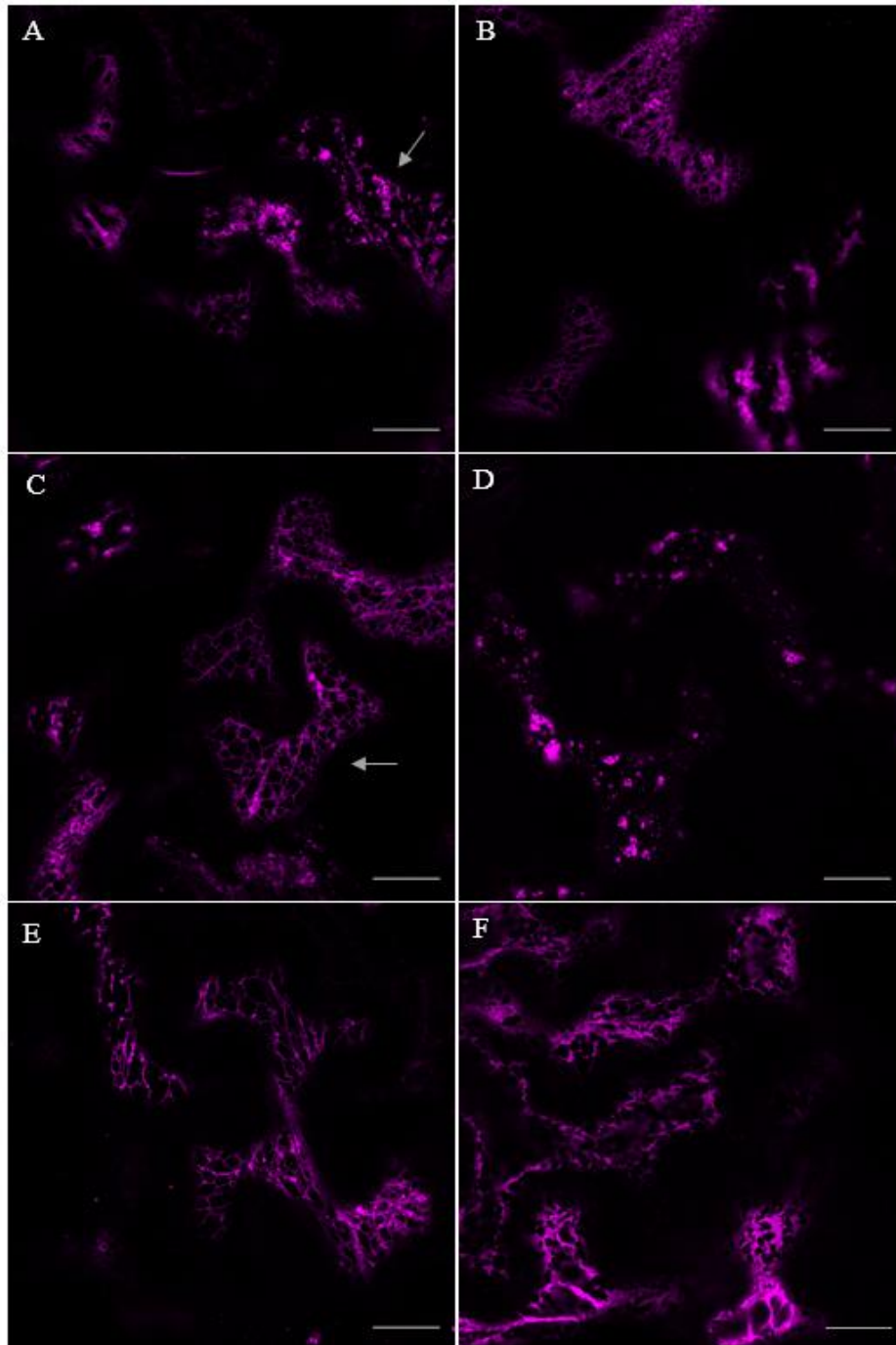


Figure 5.3. Representative images used for visual assessment of ER disruption.

Example images of (A) RFP-polypeptide + p19, (B) RFP-CAB + p19, (C) RFP-CAB-GFP + p19, (D) pmoC-RFP, (E) pmoA-RFP, (F) spBtrunc-RFP^{hdel}, at 2 days after infiltration in *Nicotiana benthamiana*. The grey arrow in (A) points to a cell with a disrupted ER, the grey arrow in (C) point to a cell with a contiguous ER. Scale bar = 20 μ m.

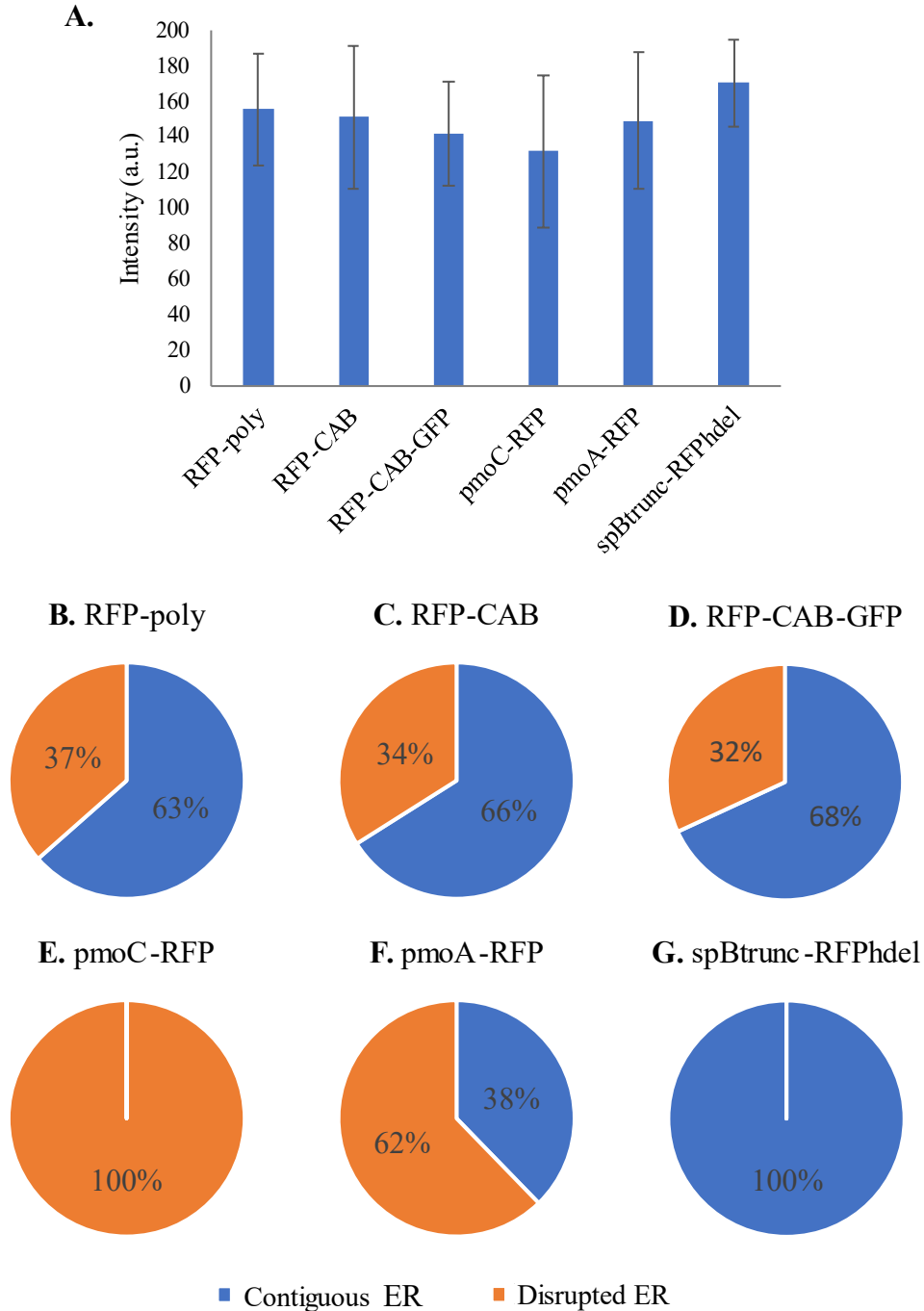


Figure 5.4. Estimation of ER disruption induced by pMMO constructs.

(A) Visual assessment of RFP expression levels of different constructs ($n=1$, technical repeats = 20 (RFP-poly), 17 (RFP-CAB), 19 (RFP-CAB-GFP), 14 (pmoC-RFP), 15 (pmoA-RFP), 16 (spBtrunc-RFP Δ del) in ROIs with similar expression levels, following transient expression in *Nicotiana benthamiana* at OD₆₀₀ 0.1 (with co-expression with p19 in the case of polypeptides). Laser power = 0.6%, gain = 600 V. The mean and standard deviation are shown. (B-G) Percentage of cells presenting either a contiguous ER or a disrupted ER phenotype within the ROIs with similar expression levels. Total number of cells counted = 63 (RFP-poly), 53 (RFP-CAB), 69 (RFP-CAB-GFP), 46 (pmoC-RFP), 53 (pmoA-RFP), 67 (spBtrunc-RFP Δ del).

5.2.2 Quantitative analysis of ER structure using AnalyzER

5.2.2.1 Experiment design

To provide more detailed analysis of the impact of polypeptide expression on ER structure and dynamics, the effect of polypeptide expression on changes to the ER was examined using the AnalyzER software. Tobacco cells expressing the luminal marker GFP-HDEL were used as a measure of wild-type ER, given that membrane markers, such as calnexin-GFP, themselves tend to induce prominent changes to ER morphology, such as an abundance of sheet formation (Runions *et al.*, 2006). Some ER metrics are also likely to differ between a luminal and a membrane protein, therefore, pMMO proteins were co-expressed with GFP-HDEL, and the metrics for the GFP-HDEL labelled lumen were extracted in each case (Figure 5.5). The pmoC subunit was not included in this analysis due to the absence of cells which presented a contiguous ER network.

Due to a difference in expression levels between RFP-CAB and the single subunits in tobacco, with RFP-CAB being highly downregulated (see section 4.2.5), polypeptide cells with higher expression levels were very infrequent, and therefore required multiple imaging sessions. Experiments were carried out in tobacco without the use of p19 to avoid potential additional modifications to ER structure caused by the large increase in GFP-HDEL which would occur as a consequence of p19. The RFP intensity levels were quantified and cells which presented similar expression levels were used (Figure 5.5 B).

The analysis was carried out using 30-frame time-series of RFP-CAB + GFP-HDEL (n= 20), spBtrunc-RFPhdel + GFP-HDEL (n= 27), and GFP-HDEL (n= 26), acquired across 7 different plants.

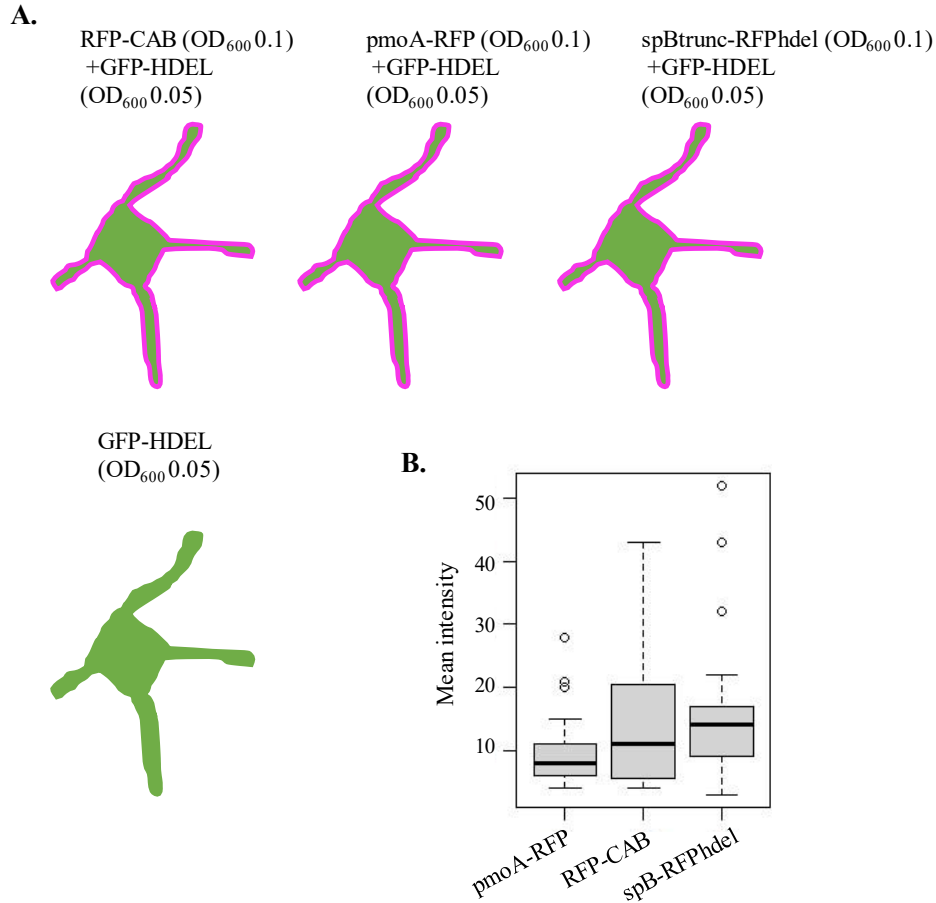


Figure 5.5. Experiment design for AnalyzER.

(A) Schematic diagram of the co-expression of pMMO proteins on the ER membrane (magenta) with GFP-HDEL in the lumen (green), and GFP-HDEL alone acting as a wild-type control. (B) Quantification of the RFP intensity level in images used for AnalyzER.

5.2.2.2 Co-expression with GFP-HDEL and skeletonization

Upon co-expression of pmoA-RFP with GFP-HDEL, the ER lumen showed a distinct change compared to GFP-HDEL alone. The lumen of GFP-HDEL alone shows a relatively homogeneous labelling of the network with a high definition of the structure of tubules and cisternae (Figure 5.6 A). Instead, the lumen in pmoA-RFP presented areas with very bright patches and areas of very faint traces of the network (Figure 5.6 B), resembling the tubule constriction phenotype induced by overexpression of some reticulon proteins. A skeleton could not be extracted consistently from images of the GFP channel with co-expression of pmoA-RFP using the software, with parts of the skeleton dropping out between frames or very small

areas of the image forming a skeleton. Therefore, pmoA-RFP was not further analyzed with the software, however, by visual inspection and by the inability to form a skeleton, it could be determined that pmoA altered the structure of the ER lumen in a greater measure than spBtrunc-RFPhdel or RFP-CAB.

There was a considerable variation in ER morphology in cells expressing RFP-CAB. The green channel of cells co-expressing RFP-CAB showed some cells in which tubules presented patches of bright green areas and some areas appeared as streams with many appressed tubules, with a low content of cisternae (Figure 5.6 F-H). However, other cells expressing RFP-CAB and all cells expressing spBtrunc-RFPhdel did not show prominent changes in the distribution of the luminal marker by visual inspection (Figure 5.6 I-N).

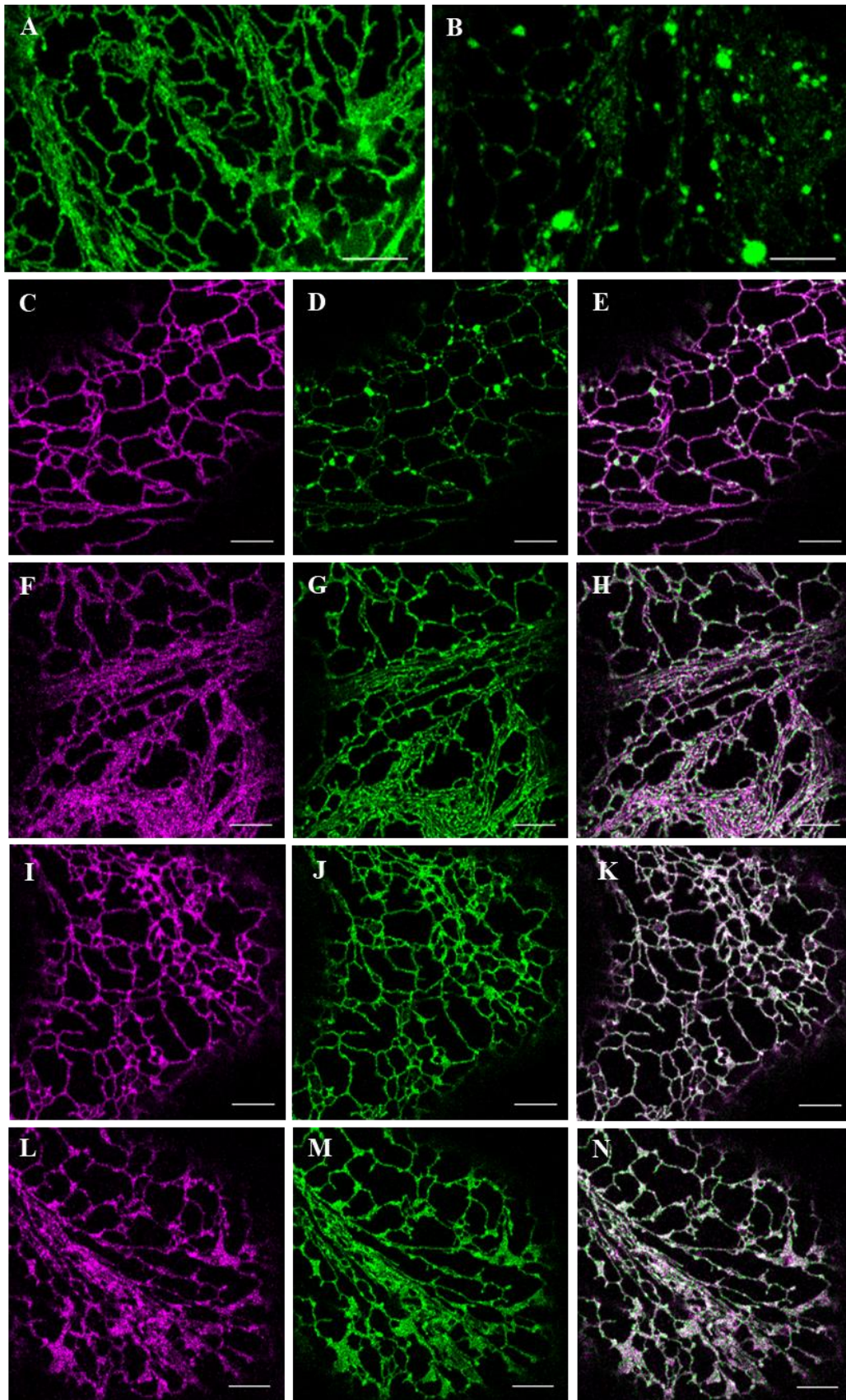


Figure 5.6. Example images used for AnalyzER.

Close view of the ER lumen in a cell expressing GFP-HDEL alone (**A**), and GFP-HDEL with pmoA-RFP (**B**). Example images of cells co-expressing pmoA-RFP with GFP-HDEL (**C-E**), RFP-CAB with GFP-HDEL showing a high content of appressed tubules (**F-H**), RFP-CAB with GFP-HDEL without a high content of appressed tubules (**I-K**), spBtrunc-RFP_{HDEL} with GFP-HDEL (**L-N**). Scale bar = 5 μm .

AnalyzER has several user-defined parameters which allow segmentation of a pixel skeleton with high fidelity. One tunable parameter is the minimum area (min A) at which a segmented region is classified as a cisterna, whereas below this value the structures are classed as nodes within the tubular network. The default min A value in the software is 0.3 μm^2 (Sparkes *et al.*, 2009), however, in the current dataset it was found that for cells expressing RFP-CAB where the lumen shows bright patches in the tubules, many of these patches were incorrectly classed as cisternae (Figure 5.7 A-B). An increase in min A to 1 μm^2 caused a reduction in the classification of these structures as cisternae (Figure 5.7 C), whilst still segmenting the ER skeleton with high fidelity in control cells expressing GFP-HDEL (Figure 5.7 D-F). The min A value of 1 μm^2 was thus adopted for the analysis.

The regions classified as cisternae included sheet-like areas, as well as streams of rapidly moving ER with a mixture of compacted sheets and tubules, and streams of very closely appressed tubules (such as that seen in Figure 5.7 A). However, cisternal texture parameters and speed metrics are expected to differ between the different kinds of cisterna-like regions (Pain *et al.*, 2019), and allow detection of differences between constructs if present.

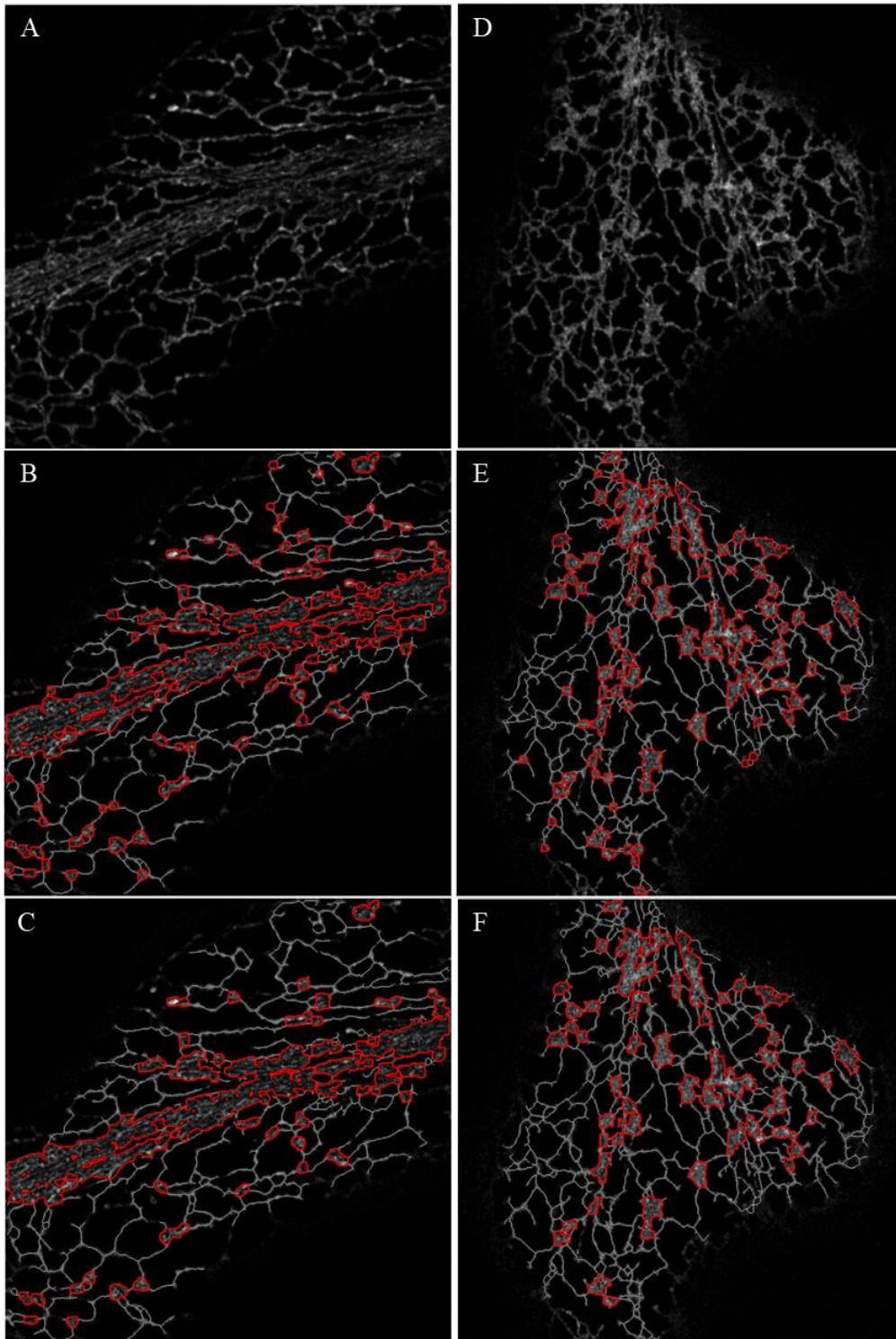


Figure 5.7. Skeletonization of the ER network.

Example of the AnalyzER processing of an RFP-CAB image following processing and filtering (**A**), skeletonization with $\min A = 0.3 \mu\text{m}^2$ (**B**), and skeletonization with $\min A = 1 \mu\text{m}^2$ (**C**). The same processing steps are shown for an example image of the GFP-HDEL control (**D-F**).

5.2.2.3 Cisternal structure, polygonal areas and tubule length

Cisternal descriptors include morphological and texture metrics. The former describes the cisternal shape, and the latter are intensity-based measures that describe the distribution pattern of the marker within the cisternae. The morphological metrics include cisternal area, circularity (the ratio of the radius calculated from the area to the radius calculated from the perimeter), elongation (the ratio of the major axis to the minor axis), and roughness (a measure of the perimeter to area ratio) (described in the AnalyzER statistics manual, (Pain *et al.*, 2019)). No major differences were observed between the constructs in these parameters, with the exception of a marginally larger cisternal area of spBtrunc-RFP_{hdel} compared with RFP-CAB and the GFP-HDEL control (Figure 5.8 A). Note that statistical analysis were not conducted for each metric separately - instead significant differences were evaluated following MANOVA analysis of multiple network parameters applying a correction for multiple comparisons (see section 5.2.4).

The cisternal texture metrics are based on gray-level co-occurrence matrices (GLCMs) and include a measure of contrast, correlation, energy and homogeneity. A completely homogeneous cisterna is expected to have contrast and energy of 0, and correlation and homogeneity of 1. In the present study, all three groups presented cisternal contrast values close to 0, i.e. reasonably even fluorescence, but spBtrunc showed a wider distribution (Figure 5.8 B). Energy was lowest for GFP-HDEL, and higher for spBtrunc and RFP-CAB, with spBtrunc again showing more variability. There were little differences in mean homogeneity but greater variance in spBtrunc. The cisternal correlation value was highest in GFP-HDEL, and decreased in RFP-CAB followed by spBtrunc (Figure 5.8 B), indicating a more uneven texture in the structures classed as cisternae in cells with pMMO proteins.

The texture metrics were initially calculated using a “distance” factor of 1, which is the distance from a pixel which is examined in four directions, where the factor is the number of radii of the smallest tubule. In cells expressing the polypeptide, several cisternae consisted of closely appressed tubules, which were discernable in both the green and the red channel (Figure 5.9

A). Some cisternae also appeared as regions with a very patchy distribution of the luminal marker, whereas the red channel presented a more homogeneous distribution (Figure 5.9 B). In order to further study the impact of these features on the network metrics, the same analysis was run using a distance of 3 to sample the environment slightly further away from each pixel. The largest difference in the trends with the increased distance parameter could be seen in the correlation, where RFP-CAB now showed the lowest correlation of all three groups (Figure 5.8 C). This reflects the greater unevenness in the luminal marker distribution seen in RFP-CAB cisternae (Figure 5.9). For all groups, the contrast values increased and the correlation values decreased, compared to using a distance of 1, revealing greater heterogeneity between a pixel and its wider neighbourhood. On the other hand, the energy and homogeneity values remained largely unaffected.

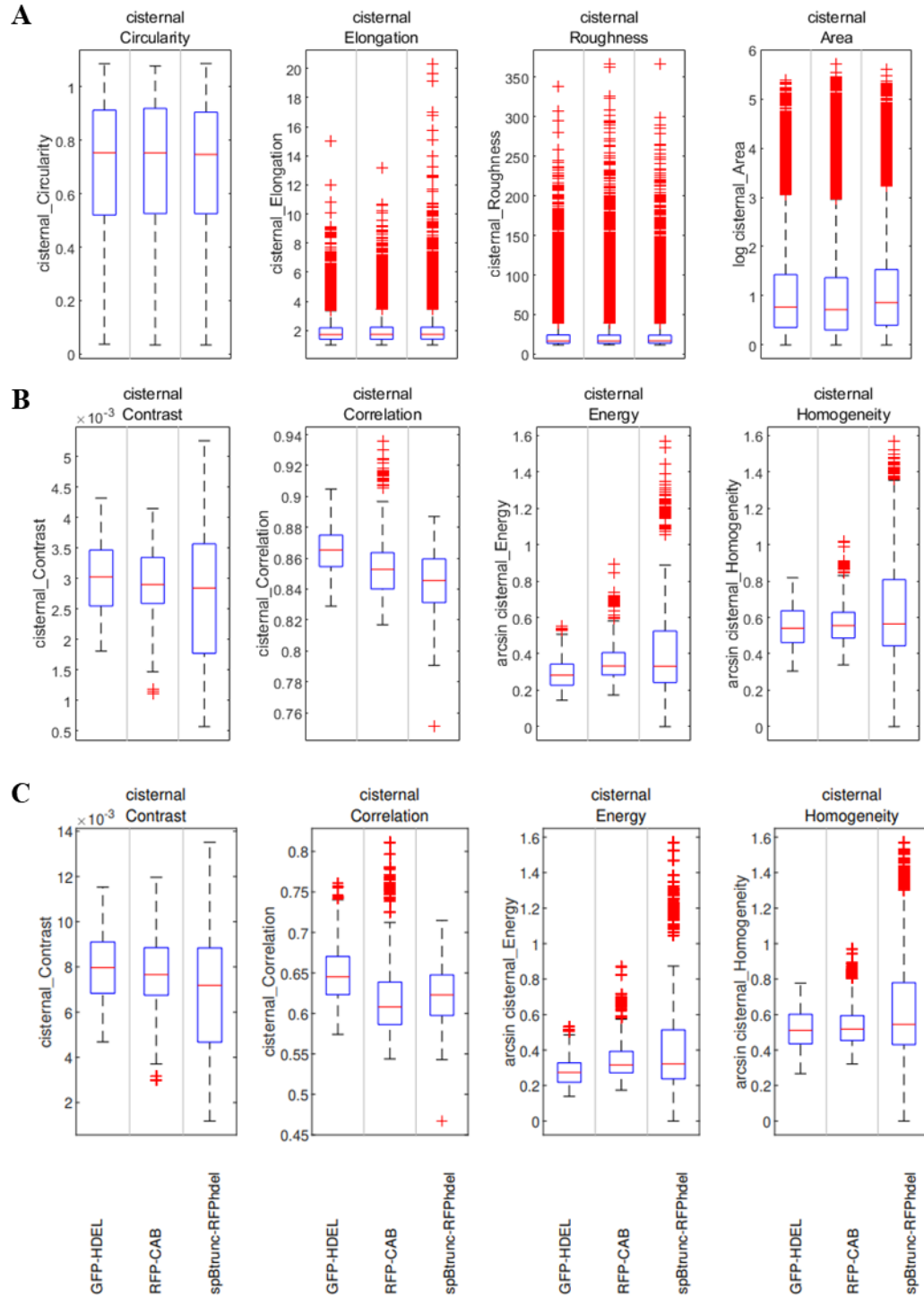


Figure 5.8. Metrics of cisternal morphology and texture.

Boxplots showing cisternal structure following AnalyzER processing of the luminal channel of cells expressing GFP-HDEL alone ($n = 26$), RFP-CAB + GFP-HDEL ($n = 20$), and spBtrunc-RFPphdel + GFP-HDEL ($n = 27$). **(A)** Boxplots of morphology parameters, **(B)** boxplots of texture parameters using distance = 1, **(C)** boxplots of texture parameters using distance = 3. Boxplots show the median (centre line), upper and lower quartiles (box limits), 1.5x interquartile range (whiskers), and outliers (red crosses).

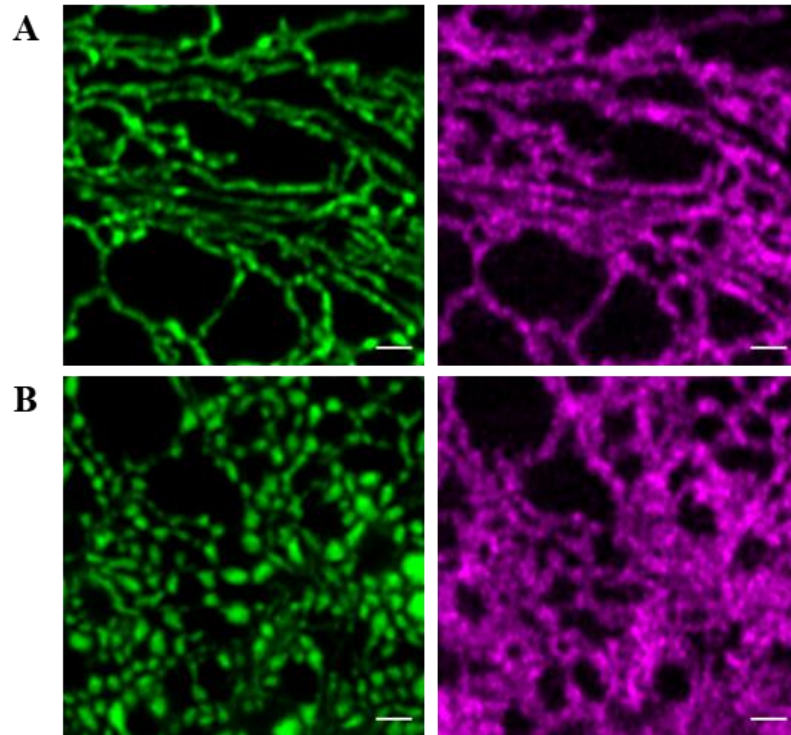


Figure 5.9. Texture of cisternae in RFP-CAB.

High magnification images showing the distribution of the luminal marker (green) found in some cases in cells expressing RFP-CAB (magenta). (A) Example of a cisternae consisting of closely appressed tubules, (B) example of a cisternae showing a patchy distribution of the luminal marker. Scale bar = 1 μm .

Other variables included to describe the ER architecture were metrics on the morphology of polygonal regions (polygon area, polygon circularity, polygon roughness, and polygon elongation) and tubule length. The groups did not appear to differ appreciably with regards to these metrics (Figure 5.10).

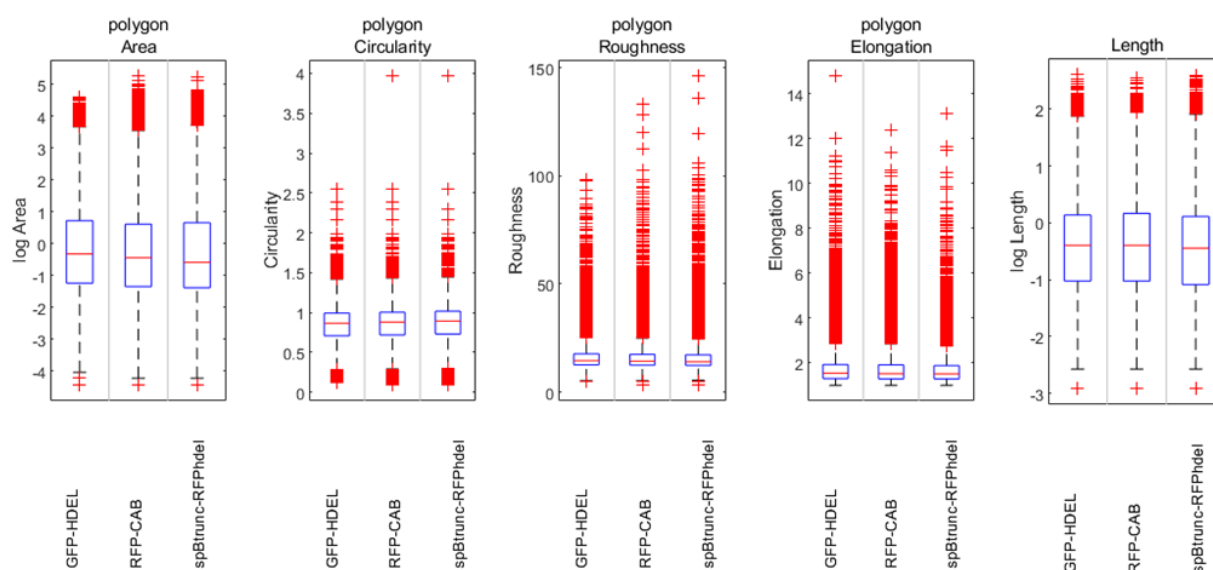


Figure 5.10. Metrics for polygon morphology and tubule length.

Boxplots showing metrics for the luminal channel on polygon area, polygon circularity, polygon roughness, polygon elongation and tubule length, for cells expressing GFP-HDEL alone ($n = 26$), RFP-CAB + GFP-HDEL ($n = 20$), and spBtrunc-RFPPhdel + GFP-HDEL ($n = 27$). Boxplots show the median (centre line), upper and lower quartiles (box limits), 1.5x interquartile range (whiskers), and outliers (red crosses).

5.2.2.4 Tubule morphology analysis

The RFP-CAB construct induced a patchy structure in the luminal marker in some cells, where brighter and dimmer areas of GFP-HDEL were observed along a tubule (Figure 5.11 A). This somewhat resembled the phenotype produced by over-expression of reticulin proteins, whereby the membrane proteins cause constriction of the ER lumen, inducing local thinning of the tubules, interspaced with discrete bulges of luminal contents (Sparkes *et al.*, 2010; Breeze *et al.*, 2016). Tubule morphology analysis using AnalyzER performed in a previous study on tobacco cells over-expressing RFP-RTN1, clearly showed the constriction phenotype by tracing the intensity values on both channels along a tubule and observing a negative correlation between the peaks and troughs in intensity between channels (Pain *et al.*, 2019). Peaks and troughs are calculated along a tubule between two nodes, with peaks identified from the pixel intensity compared to its neighbours, and a minimum separation condition. The tubule trace analysis on a high magnification image of RFP-CAB + GFP-HDEL did not show an

overall negative correlation between peaks and troughs in both channels (Figure 5.11 C): some peaks in the green channel, such as a prominent peak in the middle of the trace, coincide with a trough on the red channel, however, several other peaks and troughs in the green channel are positively correlated in parallel with the intensity changes in the red channel. Additional traces were plotted from different cells expressing RFP-CAB and show the same pattern (Supplementary Figure 5.1). Therefore, the current data suggests that RFP-CAB is not causing constriction of tubules, as the red signal is not associated with exclusion of the green. However, the peaks and troughs in the green channel are more marked than those in the red, reflecting the appearance of luminal content in bulges compared to a more homogeneous labelling of the membrane as seen by visual inspection.

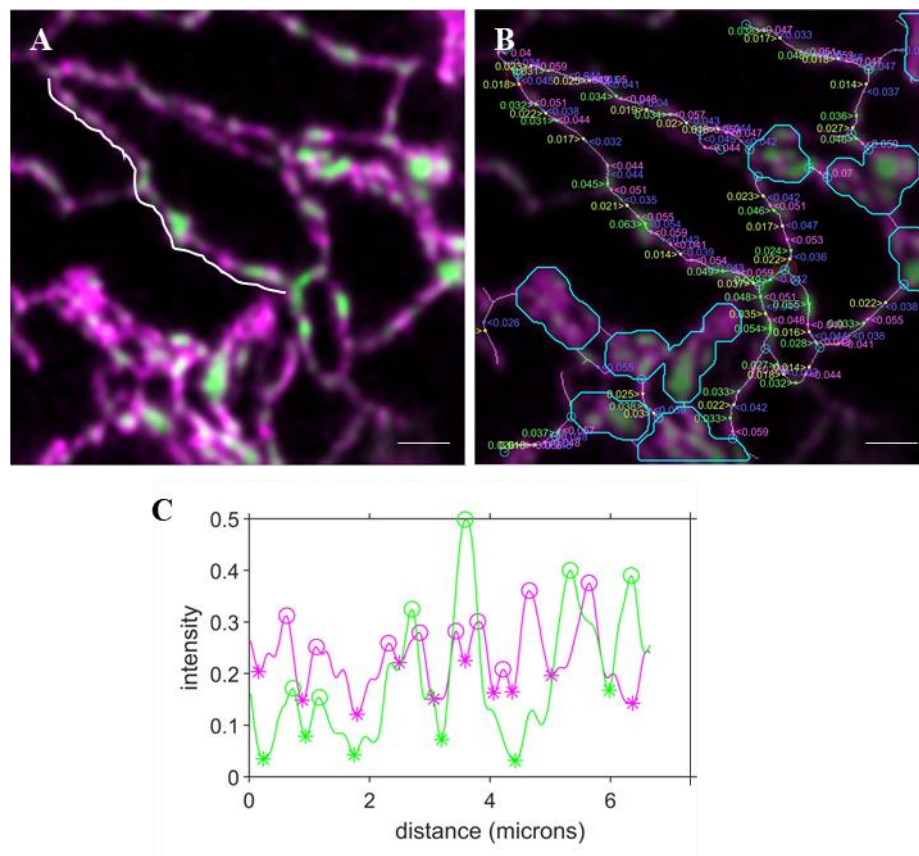


Figure 5.11. Tubule trace analysis.

(A) High magnification image of RFP-CAB + GFP-HDEL, the white line drawn adjacent to a tubule indicates the tubule which was analyzed. (B) All peaks and troughs detected in the green channel are labelled in magenta and purple respectively, and peaks and troughs on the red channel are labelled in green and yellow, displaying the estimated tubule widths. Cisternae are

circled in blue. (C) Trace plot along a tubule showing the normalized intensity values of both channels across the peaks and troughs. Scale bar = 1 μm .

5.2.3 Quantitative analysis of ER movement and persistency

The ER is a dynamic organelle which undergoes remodelling of tubules and cisternae as well faster movement along cytoplasmic streams. To study the dynamics of the plant ER in cells expressing the polypeptide the dataset used for AnalyzER consisted of 30-frame time series at 821 ms intervals. The software extracts information on speed using the Farnebäck optical flow method (Farnebäck, 2003), and provides a color-coded map where dark blue corresponds to a slower movement (Figure 5.12 A-C). Boxplots were generated representing the maximum speed in tubules (speed max), the average speed in tubules (speed global), and the maximum and average speed in cisternae, for cells expressing RFP-CAB, spBtrunc-RFP Δ DEL and GFP-HDEL (Figure 5.12 D). Based on the distributions, all four measures of speed appear similar across groups, with a slightly lower cisternal maximum speed and cisternal global speed in the case of spBtrunc. The log-transformed speeds followed a normal distribution (Supplementary Figure 5.2) and the back-transformed mean speed of cells expressing the GFP-HDEL control was 0.18 $\mu\text{m s}^{-1}$ for tubules and 0.15 $\mu\text{m s}^{-1}$ for cisternae, and the maximum speed of cells had an average of 0.34 $\mu\text{m s}^{-1}$ for tubules and 0.69 $\mu\text{m s}^{-1}$ for cisternae. These values fall within the range of previously published results for tobacco epidermal cells expressing the GFP-HDEL marker (Pain *et al.*, 2019).

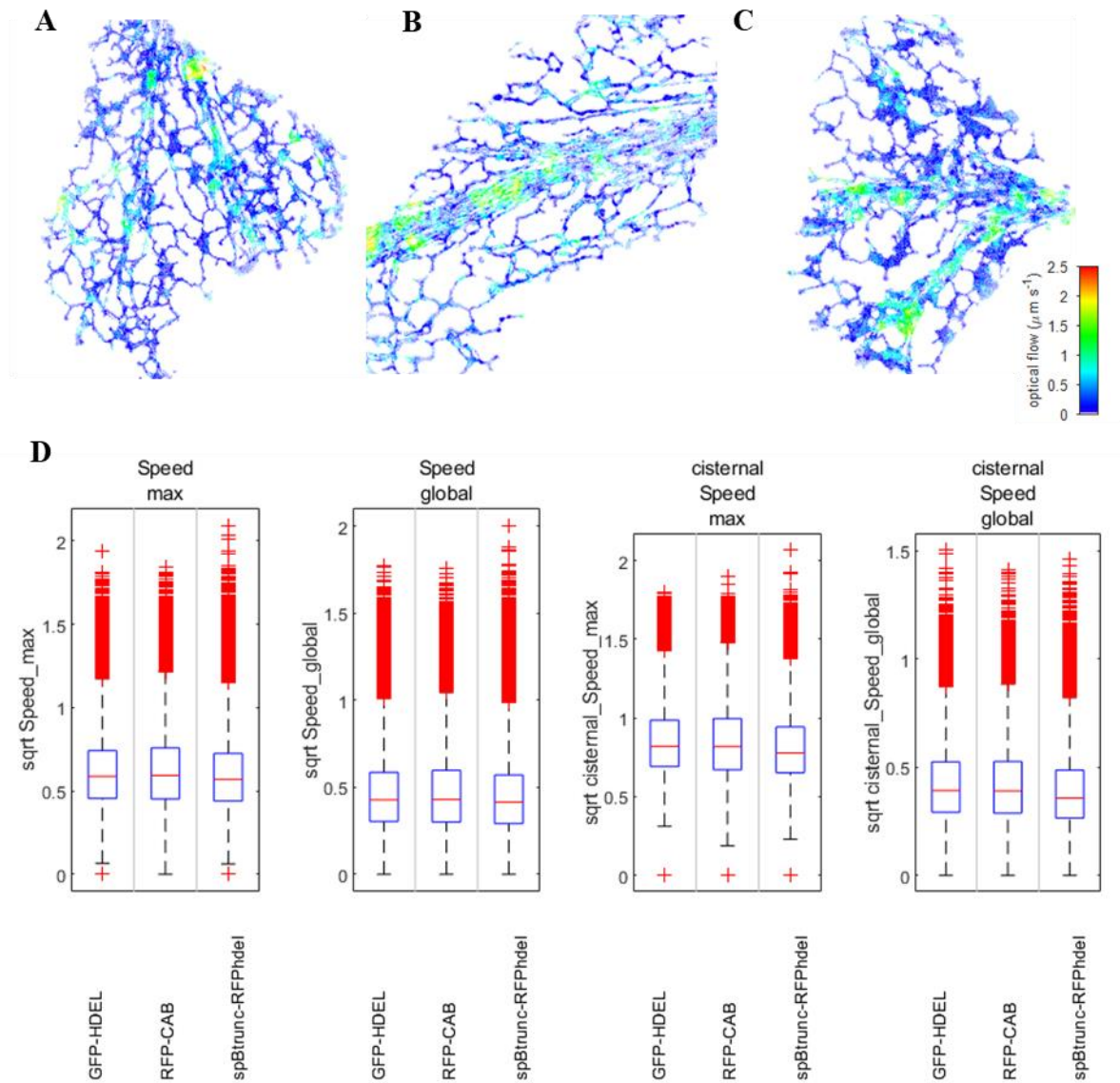


Figure 5.12. Optical flow analysis.

Example images color-coded according to speed of GFP-HDEL (A), RFP-CAB (B), and spBtrunc-RFPphdel (C). (D) Boxplots showing the maximum speed (speed max) and average speed (speed global) of tubules and cisternae, in cells expressing GFP-HDEL alone ($n = 26$), RFP-CAB + GFP-HDEL ($n = 20$), and spBtrunc-RFPphdel + GFP-HDEL ($n = 27$). Boxplots show the median (centre line), upper and lower quartiles (box limits), 1.5x interquartile range (whiskers), and outliers (red crosses).

Whereas some regions of the ER undergo rapid movement, other areas in tubules and cisternae are less mobile and remain more persistent over time. Particularly persistent regions are often nodes associated with points of contact between the ER and the plasma membrane (Sparkes *et al.*, 2009; Griffing *et al.*, 2016). Persistency analysis was carried out to further characterize the

ER dynamics, using a lag period of 7 frames (approximately 5.7 s) above which the pixels were classified as persistent. Example images show the persistency mapping using a color scale with bright shades identifying regions of high persistency in cisternae (magenta) and tubules (green) (Figure 5.13 A-C). Boxplots show that the persistency in tubules was similar across groups, whereas the mean cisternal persistency appeared to be marginally higher in spBtrunc-RFPhdel (Figure 5.13 D).

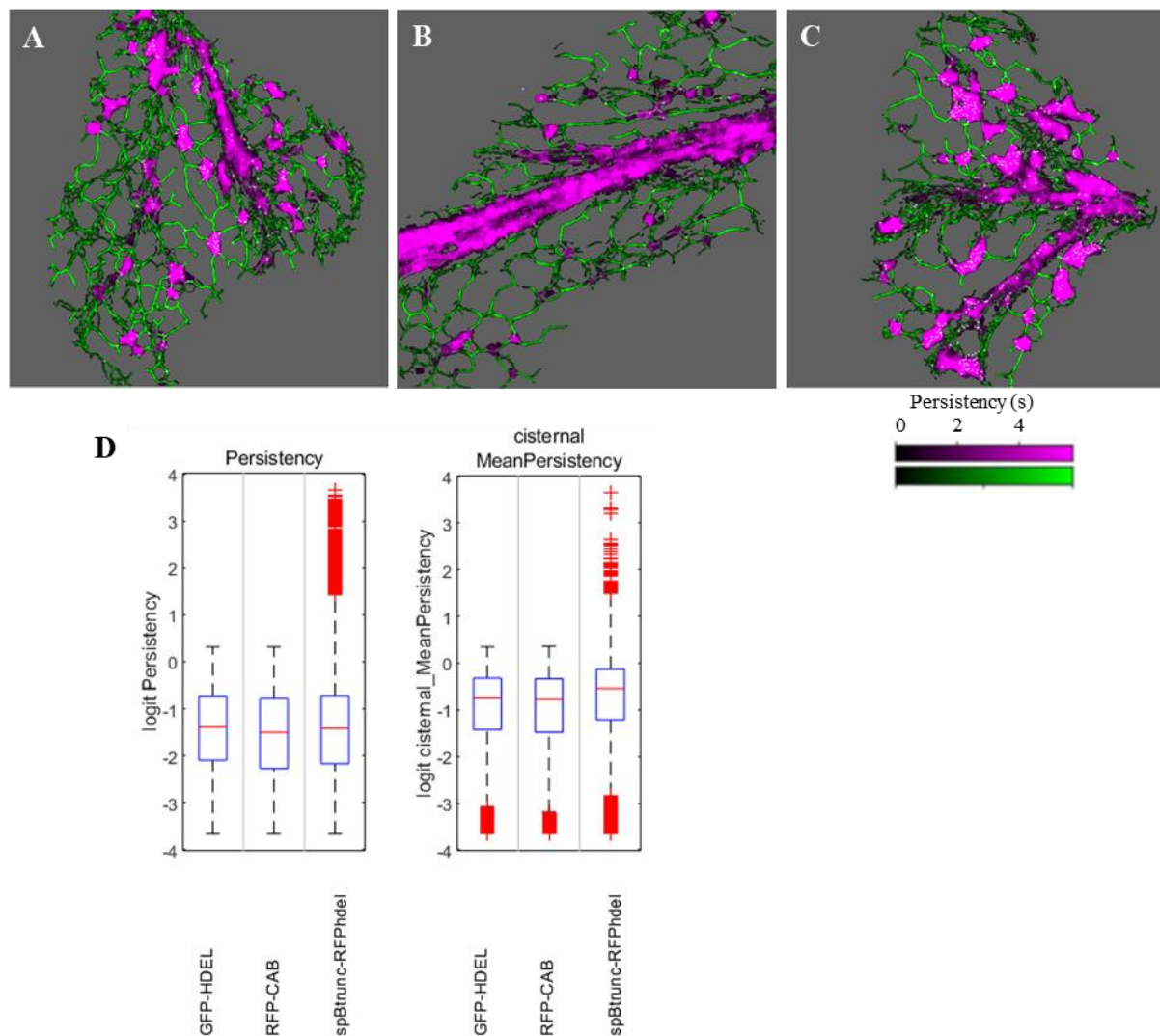


Figure 5.13. Persistency analysis.

Example images showing the persistency of static elements in cisternae (magenta) and tubules (green) in GFP-HDEL (A), RFP-CAB (B), and spBtrunc-RFPhdel (C). (D) Boxplots showing the mean persistency of tubules (Persistency) and cisternae (cisternal MeanPersistency) in cells expressing RFP-CAB + GFP-HDEL (n = 20), spBtrunc-RFPhdel + GFP-HDEL (n = 27), and

GFP-HDEL alone (n = 26). Boxplots show the median (centre line), upper and lower quartiles (box limits), 1.5x interquartile range (whiskers), and outliers (red crosses).

5.2.4 Network comparisons using multiple parameters

The ER network is a complex structure which can be characterized by multiple parameters as described above, including the morphology of tubules, cisternae, and polygonal areas, as well as its movement and regions of higher persistency. In order to analyze the impact of the expression of the polypeptide upon ER architecture as a whole, a multivariate analysis of variance (MANOVA) was performed, combining 19 metrics of ER structure and dynamics (Pain *et al.*, 2019). Transformations were applied to the variables where necessary to achieve or approximate normal distributions. The MANOVA showed that significant differences existed in the ER architecture between the groups (Pillai's trace, $F(3.227)$, $p = 1.5 \times 10^{-6}$; Roy's largest root, $F(5.7947)$, $p = 2.39 \times 10^{-7}$). All three groups were separated using two canonical variables, with spBtrunc-RFPhdel separating on the first canonical variable, and RFP-CAB separating from GFP-HDEL on the second canonical variable (Figure 5.14 A). The dendrogram plotting the group's multivariate means shows a very close similarity between the network of cells expressing RFP-CAB and the GFP-HDEL control, and a larger separation of spBtrunc-RFPhdel with respect to the other two groups (Figure 5.14 B). The multivariate analysis was also performed using the distance = 3 parameter for cisternal texture, obtaining the same grouping of the constructs (Supplementary Figure 5.3).

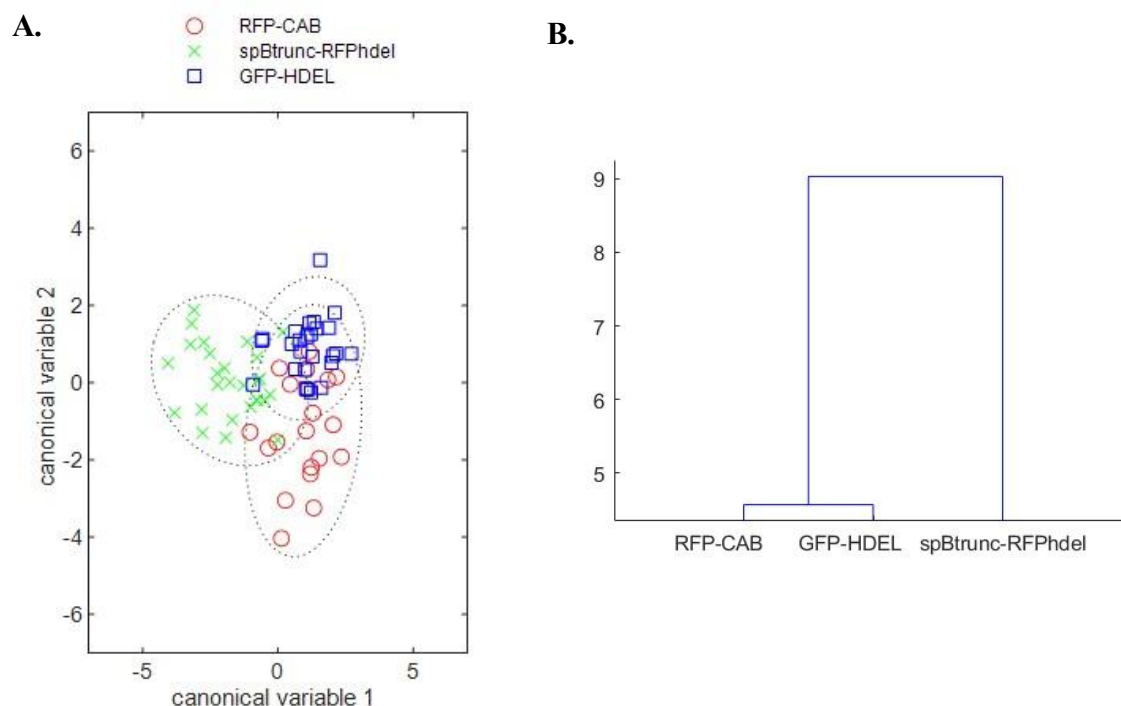


Figure 5.14. Multivariate analysis of ER structure and dynamics.

(A) MANOVA performed using 19 dependent variables characterizing the ER network, on cells expressing RFP-CAB + GFP-HDEL ($n = 20$), spBtrunc-RFPhdel + GFP-HDEL ($n = 27$), and GFP-HDEL alone ($n = 26$). Dotted lines represent the 95 % confidence limit. (B) Dendrogram showing the clustering of the three different groups based on their multivariate means.

An ANOVA was performed using the same dataset to test which variables presented significant differences in the means between groups. Significant differences were detected in cisternal correlation, cisternal maximum speed, cisternal global speed, and cisternal mean persistency (Table 5.1). Post-hoc comparisons were carried out to determine specific differences between the groups. The Tukey HSD method was used to correct for multiple comparisons and control the type I error. The Tukey comparisons showed a significant difference in cisternal correlation between RFP-CAB and GFP-HDEL, and between spBtrunc-RFPhdel and GFP-HDEL. Differences were also detected between spBtrunc-RFPhdel and GFP-HDEL in cisternal maximum speed, and between spBtrunc-RFPhdel and both other groups in cisternal mean persistency, with spBtrunc-RFPhdel presenting a lower maximum speed and higher persistency of cisternae. Taken together this would indicate that considering the combined effect of the

variables measured, the RFP-CAB polypeptide does not induce major changes to the ER architecture and dynamics. In the case of spBtrunc, the significant multivariate response arises from the combination of several parameters.

Table 5.1. ANOVA results and pairwise post-hoc comparisons.

ANOVA carried out with the same dataset used for MANOVA and boxplots, using distance = 1 for cisternal correlation. The following variables were transformed to approximate a normal distribution: tubule length (log), speed max and speed global (square root), persistency (logit), cisternal energy and homogeneity (arcsin), cisternal area and polygon area (log). Tukey HSD comparisons are shown as : * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

Variable	F	p-value	RFP-CAB/ GFP-HDEL	spB-RFPhdel/ GFP-HDEL	RFP-CAB/ spB-RFPhdel
Tubule length	0.4124	0.664			
Tubule speed max	1.998	0.143			
Tubule speed global	1.37	0.261			
Tubule persistency	1.72	0.187			
Cisternal contrast	0.1929	0.825			
Cisternal correlation	9.767	0.000182	***	**	
Cisternal energy	0.3012	0.741			
Cisternal homogeneity	0.05686	0.945			
Cisternal speed max	3.769	0.0279		*	
Cisternal speed global	3.309	0.0423			
Cisternal persistency	7.455	0.00116		*	**
Cisternal circularity	0.1594	0.853			
Cisternal elongation	0.4431	0.644			
Cisternal Roughness	0.6224	0.54			
Cisternal area	1.828	0.168			
Polygon area	0.1901	0.827			
Polygon circularity	2.67	0.0763			
Polygon Roughness	1.668	0.196			

Polygon elongation	2.663	0.0768			
-----------------------	-------	--------	--	--	--

5.3 Discussion

5.3.1 Network morphology and texture metrics

The ER is an extensive membranous compartment which has the potential to harbour large quantities of membrane proteins. However, over-expression of membrane proteins has often been reported to induce substantial changes to ER morphology and dynamics (Runions *et al.*, 2006; Sparkes *et al.*, 2010; Kriechbaumer *et al.*, 2018; Lazareva *et al.*, 2021). Although a specific link between ER structure and function has not been established, it is likely that severe changes to ER morphology or dynamic may impact upon cellular functions amongst which the ER plays a role. Therefore, having localized all three pMMO subunits to the ER and co-expressed them through polypeptides, their impact upon ER structure and movement was studied.

Recombinant expression of polypeptides containing all three pMMO subunits (RFP-polypeptide, RFP-CAB, RFP-CAB-GFP) at high levels by co-expression with p19, resulted in over 60% of cells presenting a contiguous ER network (Figures 5.3 and 5.4), meaning that cells are able to express the pMMO polypeptides without severe damage to the ER. Polypeptide expression appeared to be substantially less disruptive than the production of pmoC alone - which resulted in disruption of the ER in all cells - but more disruptive than the production of spBtrunc alone, which did not result in any substantial perturbation.

Co-expression of the pMMO single subunits and RFP-polypeptide with GFP-HDEL allowed analysis of the effects of these proteins on the ER and to perform a quantitative analysis using AnalyzER. Surprisingly, the pmoA subunit induced a phenotype which resembled the characteristics of reticulon over-expression, or the “beads on a string” pattern of the lumen. Constriction produced by reticulon proteins is reported to be mediated by their conserved reticulon homology domain, composed of four transmembrane domains oriented in a W shape, and specifically by an amphipathic helix within this region (Sparkes *et al.*, 2010; Breeze *et al.*,

2016). A viral movement protein localized to the plant ER was also reported to induce tubule constrictions, and hypothesized to do so through a W-shaped topology resembling the reticulon topology, with both termini facing the lumen (Lazareva *et al.*, 2021). The pmoA subunit was found to insert in the plant ER with both termini facing the cytosol (see 3.2.5), and possesses six transmembrane domains. Its topology could provide a mechanism for membrane curvature by insertion in a wedge-like manner into the membrane similarly to reticulons. However, other characteristics such as the length of the TMDs, their insertion at an angle, or oligomerization properties, may also be the motive of the observed phenotype. The phenotype produced by pmoA with a great unevenness in the distribution of the luminal markers resulted in the inability to skeletonize the lumen, and therefore the subunit was excluded from AnalyzER processing. Co-expression of RFP-polypeptide with GFP-HDEL revealed by visual inspection that approximately 50% of cells contained a patchy texture of the lumen, with brighter and dimmer regions across tubules, as well as a high proportion of streams of tightly appressed tubules (Figures 5.6 and 5.9). Regions of closely appressed tubules are classed as cisternae using AnalyzER, and their presence, as well as the patchy cisternae, were reflected in a significant decrease in the cisternal correlation of RFP-CAB with respect to the GFP-HDEL control (Figure 5.8 and Table 5.1), indicative of less homogeneous cisternae. A higher cisternal energy was also observed for RFP-CAB with respect to the control but no significant differences were found in post-hoc comparisons. No significant differences in network morphology parameters, such as the area, circularity, elongation and roughness of cisternae and polygonal regions, and the tubule length, were found between the polypeptide and the control.

5.3.2 The polypeptide does not cause constriction in tubules

In order to further analyze the patchy pattern found in tubules of some cells expressing RFP-CAB, trace graphs were plotted along tubules showing the intensity in the red and green channels. In the case of a reticulon constriction phenotype, a negative correlation is seen between the peaks and troughs in intensity in both channels, i.e. peaks in the red channel will

coincide with troughs on the green channel and vice versa (Figure 5.15, Pain *et al.*, 2019). In cells expressing RFP-CAB, a clear negative correlation was not observed, instead three different scenarios were observed: some peaks in the red channel coincided with peaks in the green channel, indicating that an accumulation of RFP-CAB was not causing exclusion of GFP-HDEL, in other cases peaks in the green channel were accompanied by a constant intensity level in the red channel, and finally some peaks and troughs did show a negative correlation between channels (Figure 5.11 and Supplementary Figure 5.1). These results would suggest that the polypeptide is not causing constriction of the luminal content in tubules. This is supported by the visual inspection of the luminal marker in example cells, where a more homogeneous labelling of the lumen is found in cells expressing RFP-CAB than those expressing pmoA-RFP, with the latter resembling the reticulon phenotype to a greater extent (Figure 5.6).

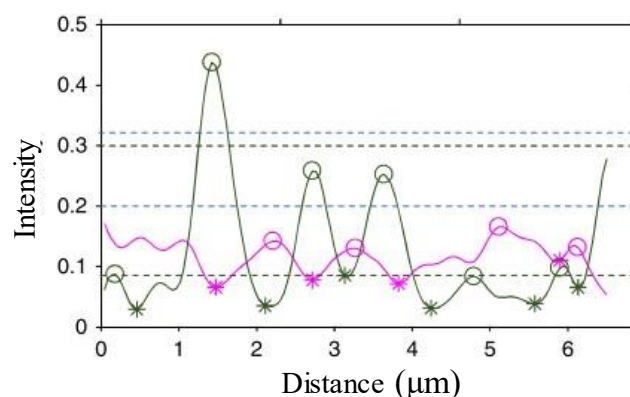


Figure 5.15. Trace diagram of RTN1 showing a typical constriction phenotype.

The normalized intensities across a tubule expressing GFP-HDEL and RFP-RTN1 are plotted. A negative correlation in the intensity between the channels can be observed, where peaks in the green channel clearly coincide with troughs on the red channel, and vice versa. Taken from (Pain *et al.*, 2019).

Perhaps the interactions between the subunits in the polypeptide are responsible for the irregular pattern of the lumen. The subunits can interact to form a heterotrimer and also to form a nonamer in methanotroph bacteria. In tobacco cells, pairwise interactions between the subunits were found as well as homotypic interactions of all three subunits. Potentially,

different proportions of quaternary structures are formed in plants, especially if the subunit stoichiometry produced from the polypeptide is not 1:1:1, and possibly certain oligomeric structures could be causing an irregular spacing of the lumen content and the different relationships between the intensity peaks found in the trace plots. For example, a higher proportion of homotypic interactions of pmoA, or pmoA with pmoC, containing an elevated content of transmembrane domains, may cause some level of constriction and produce the peaks such as those highlighted in Figure 5.15 D. Instead, a change in the proportion towards other quaternary structures may introduce slight bulges in the membrane causing peaks in the lumen (green channel) without a trough in the red channel, as seen in Figure 5.15 B and C.

5.3.3 Impact of pMMO proteins on dynamics and on the multivariate characterization of the network

The analysis of speed and persistency showed no significant differences between the polypeptide and the GFP-HDEL control. However, the ER-targeted pmoB subunit (spBtrunc) presented a significantly lower cisternal maximum speed than GFP-HDEL, and a lower cisternal average speed although this difference was not significant (Table 5.1). SpBtrunc also gave a higher mean cisternal persistency. These differences could be due to spBtrunc presenting a slightly higher cisternal area, although again, a significant difference with post-hoc comparisons was not detected for this metric. The speeds calculated in this study are slightly lower than those calculated with AnalyzER for tobacco cells expressing GFP-HDEL previously (Pain *et al.*, 2019). For tubules the average and maximum speeds were $0.18 \mu\text{m s}^{-1}$ and $0.34 \mu\text{m s}^{-1}$ respectively, and the previously reported speeds were $0.26 \mu\text{m s}^{-1}$ and $0.52 \mu\text{m s}^{-1}$; for cisternae the average and maximum speeds were $0.15 \mu\text{m s}^{-1}$ and $0.69 \mu\text{m s}^{-1}$ respectively, and the previously found speeds were $0.21 \mu\text{m s}^{-1}$ and $0.95 \mu\text{m s}^{-1}$. The ER structure and its dynamics are highly sensitive to environmental changes, therefore it is likely that these observed differences in speed could respond to systematic differences in the data

taking between the two datasets, such as the temperature of the room, or greenhouse conditions of the plants, amongst others.

A MANOVA analysis incorporating 19 metrics of ER structure and dynamics revealed the separation of all three groups in two canonical variables (Figure 5.14 A). The accompanying dendrogram showed a higher similarity between the polypeptide and the GFP-HDEL control, with spBtrunc separating from both groups (Figure 5.14 B). The MANOVA shows the combined effects of several metrics, providing a good approximation to characterizing the ER architecture as a whole. Post-hoc comparisons revealed only one metric presenting significant differences between the polypeptide and the control, and three metrics showing significant differences between spBtrunc and the control. This differs markedly from the results obtained in a previous study where a much higher number of variables were found to differ between the control and the application of drug treatments, heat shock or over-expression of LNP2 (Pain *et al.*, 2019). This indicates that taking into account the metrics considered and the parameters used for skeletonization, the ER architecture would not seem to be considerably altered with the expression of the pMMO polypeptide or with spBtrunc. Moreover, expressing the polypeptide produced less disruption or alteration of the ER network than expressing the pmoC or pmoA subunits individually. As a conclusion, the ER appears to be a suitable membrane compartment in which to recombinantly express pMMO, without producing considerable disruption of the network or changes to the ER structure and dynamics, despite the high content of transmembrane domains of the complex and its oligomerization capacities.

6 Towards assays for pMMO in plants with stable expression of pMMO

6.1 Introduction

6.1.1 Measurement of the enzymatic activity of pMMO

The pMMO enzyme complex converts methane to methanol as a first step in the oxidation of methane for its use as a carbon and energy source in methanotrophs. The reaction uses oxygen and two reducing equivalents, thought to derive from ubiquinol produced by NADH:quinone oxidoreductase (NDH-2) in methanotroph membranes (Sirajuddin and Rosenzweig, 2015). Research on pMMO activity has been predominantly based on methanotrophs due to very limited success of heterologous expression of the complex in other systems (Koo and Rosenzweig, 2021).

To detect enzymatic activity of pMMO, the rate of methane consumption can be measured. For example, one study carried out whole-cell enzyme assays of *Methylosinus trichosporium* OB3b by adding known concentrations of methane to sealed vials containing the cells, and measuring the decrease in methane concentration in the headspace using gas chromatography over an 8 hour time course (Lontoh and Semrau, 1998). However, studies measuring methane breakdown are scarce, instead the vast majority of work on pMMO determines activity via a propylene epoxidation assay, making use of the enzyme's ability to oxidize propylene gas into propylene oxide, a volatile liquid which is not produced in the cell, which can therefore be easily detected. During the propylene epoxidation assay, samples containing pMMO are placed in glass vials and sealed, following which air is removed and propylene gas is injected using a gas-tight syringe (Kitmitto *et al.*, 2005). The samples are shaken at the appropriate temperature (usually 30°C for *Methylosinus trichosporium* OB3b or 45 °C for *Methylococcus capsulatus* (Bath)),

following which an aliquot of liquid sample is injected into a gas chromatograph, and the results compared against propylene epoxide standards (Kitmitto *et al.*, 2005). pMMO activity has been assayed in methanotroph whole-cells, membrane preparations, and solubilized or purified pMMO. Nevertheless, obtaining active pMMO in the lab has proven difficult, presumably due to the enzyme being a multi subunit integral membrane complex and containing metal co-factors in its active site. Furthermore, different variables for the isolation of the enzyme and activity assays have been optimized differently by different research groups.

In terms of supplying an electron donor for the reaction, early studies discovered that NADH can act as a reducing equivalent for assays in whole-cells and in membranes, but not for solubilized pMMO (Shiemke *et al.*, 1995). Instead, the addition of duroquinol, a synthetic compound obtained from the oxidation of duroquinone, yielded activity in detergent-solubilized pMMO, as well as activity in whole-cell samples and membranes (Shiemke *et al.*, 1995), and has been thereafter routinely added in pMMO assays as a reductant. It is thus believed that pMMO uses quinols as a reducing agent in the native methanotrophs, such as ubiquinol, which is produced by a type 2 NADH dehydrogenase (NDH-2) present in methanotroph membranes (Sirajuddin and Rosenzweig, 2015), explaining why NADH functions as an electron donor only when membrane fractions or whole-cells are assayed.

Another important condition for assaying pMMO activity is the presence of copper. Early on it was observed that high levels of copper in the methanotroph membranes correlated with high enzymatic activity *in vitro*, presumably to stabilize and activate the enzyme due to the copper present in the active site (Nguyen *et al.*, 1994). Nguyen *et al.* (1994) purified membrane fractions from *Methylococcus capsulatus* (Bath) using 0-20 μM CuSO_4 in the growth medium, and found an enhancement in the specific activity of the membranes with increases in the copper levels. Moreover, the addition of supplemental copper to the assay buffer (150 μM CuSO_4) produced higher specific activities in membranes from cells grown between 0-10 μM CuSO_4 (Nguyen *et al.*, 1994). Likewise, other groups have used varied copper concentrations both in the growth medium and in other stages during purification or as supplementation in the

assays. For example, Kitmitto *et al.* (2005) assayed fractions of membrane-bound, solubilized and purified pMMO, using a final concentration of 40 μM CuSO_4 in the growth medium, and supplemented the propylene epoxidation assay reactions with 50-20 μM CuSO_4 for “optimal” activity (Kitmitto *et al.*, 2005). Another group working with membrane fractions used 10-20 μM CuSO_4 in the growth medium and supplemented cells prior to lysis with 100 μM CuCl_2 (Takeguchi *et al.*, 1998). Studies from the Rosenzweig group which produced crystal structures of pMMO, used 50 μM CuSO_4 in the growth medium, and a 500 μM CuSO_4 supplement before lysis and purification from *Methylosinus trichosporium* OB3b cells (Hakemian *et al.*, 2008), or 50 μM CuSO_4 in the growth medium without an additional supplement before lysis from *Methylocystis* species Strain M (Smith *et al.*, 2011).

Besides copper, some studies have also found an increase in pMMO enzyme activity with the addition of iron to the growth medium of *Methylococcus capsulatus* (Bath) (Zahn and Dispirito, 1996), and supplementation with iron to the membrane fraction of *Methylosinus trichosporium* OB3b (Takeguchi *et al.*, 1998). However, the effect of adding iron is not clear, with other studies strongly stating that enzymatic activity levels are only dependent on copper addition (Sirajuddin *et al.*, 2014).

Finally, a single study has reported the heterologous expression of pMMO, which was expressed in *Rhodococcus erythropolis* bacteria (Gou *et al.*, 2006). The authors optimized the growth conditions of the recombinant strain, achieving the highest cell concentration and enzyme activity with 7.5 μM Cu^{2+} and 1.8 μM Zn^{2+} . Activity was measured both via the propylene epoxidation assay and by measuring methanol production from methane. However, the specific activity of the recombinant was extremely low.

6.1.2 Objectives

The aim of this chapter was to test the enzymatic activity of pMMO expressed in plant tissue. Having achieved pMMO production from transient expression in tobacco (see Chapter 5), the creation of stable plants expressing the pMMO polypeptide was carried out to provide a

consistent source of material, and different strategies were evaluated for the set-up of enzyme assays. Both stable Arabidopsis and rice plants were generated. Arabidopsis being a useful model system for optimization and rice as the target crop for engineering.

The activity of pMMO has not been assayed in plant samples before, therefore as a first attempt to detect functional enzyme, the propylene epoxidation assay was chosen, as this is the most widely documented method and possibly more sensitive than measuring methane uptake. It was also decided to carry out propylene epoxidation *in vitro* using plant lysates, as this offers certain advantages with respect to *in vivo* assays using whole seedlings, including more direct contact between the enzyme in plant lysates and the propylene gas, higher concentration of the enzyme in plant lysates compared to whole seedlings, technical ease to extract the propylene oxide from the sample, and ease to supplement the assay with co-factors and other additives (such as duroquinol and copper).

These advantages should offer higher sensitivity for the detection of the propylene oxide, and provide flexibility to modify assay conditions. *In vivo* assays using whole seedlings and measuring methane uptake were considered the next stage in evaluating functionality of pMMO in plants, contingent on results obtained with *in vitro* assays.

The specific objectives were:

- 1) To produce stable Arabidopsis plants containing the pMMO polypeptide (expressing all three ER-localized subunits separated by self-cleaving peptides)
- 2) To analyze the effect of amino acid additions to the enzyme complex derived from excision of the self-cleaving peptides which could affect enzyme function
- 3) To create an alternative polypeptide without residual amino acid footprints and to transiently express this construct in tobacco leaves
- 4) To perform *in vitro* propylene epoxidation assays on samples derived from transient expression and stable plants
- 5) To produce stable rice plants containing the self-cleaving polypeptide

Objective 4 was to be carried out in collaboration with Professor Tom Smith at Sheffield Hallam University, due to the group's expertise in methane monooxygenase enzymes and the presence of the equipment required for the propylene epoxidation assays. However, due to COVID-19 restrictions, as well as technical issues arising during the first enzyme trials, objective 5 could not be completed. The plant samples were prepared for testing, however, results from the assays are not yet available. The experiments performed up to that point are discussed in this chapter as well as the rationale for the design of the enzyme assays.

6.2 Results

6.2.1 Stable Arabidopsis plants expressing pMMO

Different strategies were undertaken to obtain samples for the assessment of enzyme activity of recombinant pMMO in plants. Firstly, stable Arabidopsis plants were created using the floral dip method, with the RFP-polypeptide and the RFP-CAB constructs, which contain all three pMMO subunits separated by 2A self-cleaving peptides. Arabidopsis was chosen as a model plant as an alternative to tobacco to create stable transgenics, due to the longer time frame required to create stable tobacco plants through the regeneration of roots and shoots from leaf explants (Sparkes *et al.*, 2006). Furthermore, transformation with RFP-polypeptide would allow rapid screening of seedlings through the C-terminal clover reporter. Transformation was also carried out with RFP-CAB to test an alternative to the full polypeptide which would not have any additions to the C-terminus of spBtrunc deriving from the final self-cleaving peptide, in case this addition somewhat hindered functionality, whilst still containing an RFP fluorescent tag to screen the transformants.

One putative transformant was recovered from selection on kanamycin (50 mgL⁻¹) for RFP-polypeptide, and one for RFP-CAB (Figure 6.1). The seedlings were transferred to soil after 10 days on plates and grew until flowering stage with no apparent growth alterations or distinguishable phenotypes. Expression of the pMMO genes was assessed by confocal microscopy on a leaf section of mature T1 plants.

Expression of the clover fluorophore could be seen in all cells, with localization in the cytoplasm, as expected according to the results of transient expression in tobacco (Figure 6.2 A and B), indicating that the plant had undergone transformation with the polypeptide. However, no expression of RFP-pmoC could be observed (Figure 6.2 C and D). This suggested that in stably transformed Arabidopsis the pmoC subunit (and presumably pmoA and spBtrunc)

is silenced or degraded. Similarly, no expression of RFP-pmoC was observed in the RFP-CAB transformant at flowering stage (Figure 6.2 E).

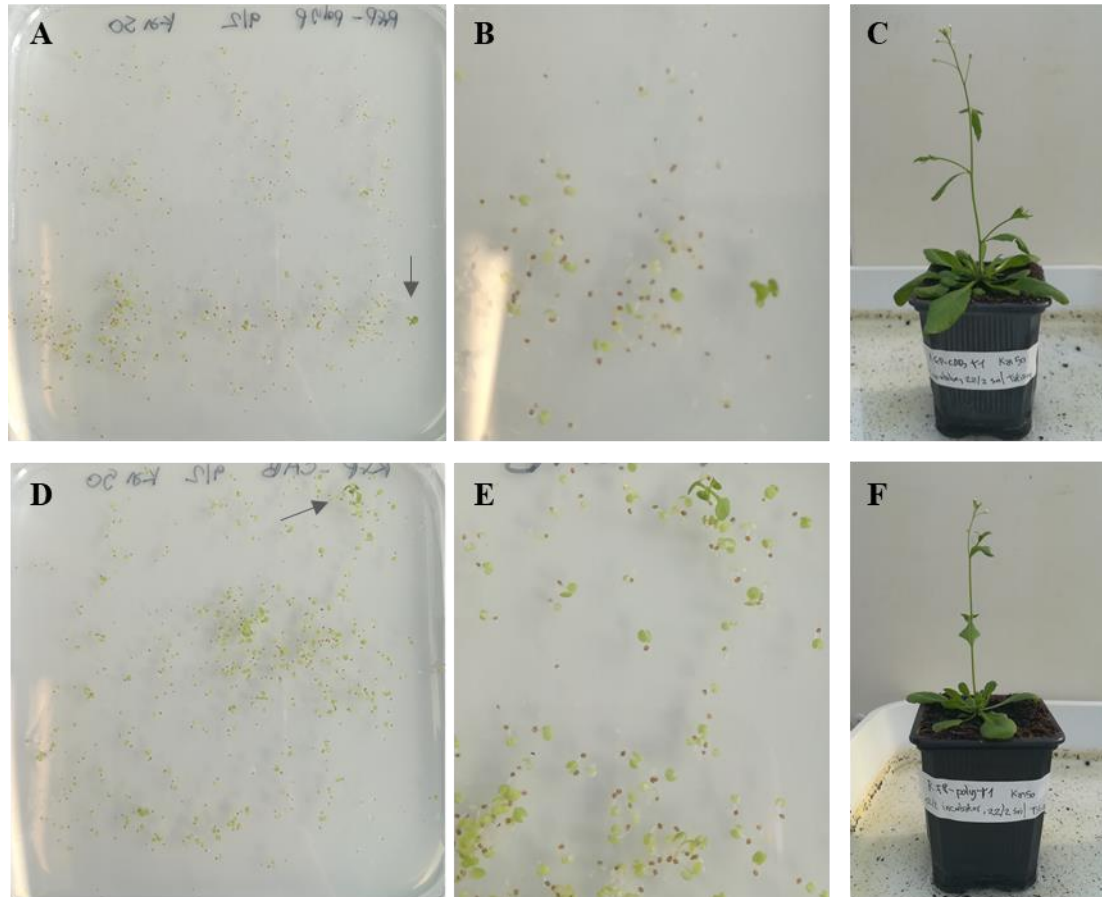


Figure 6.1. Selection of stable Arabidopsis.

MS plates with kanamycin (50 mgL⁻¹) of T1 seeds following floral dipping of RFP-polypeptide with an arrow indicating a putative transformant after 10 days (A) and closer view of the transformant (B). Selection plate with kanamycin (50 mgL⁻¹) of RFP-CAB with arrow pointing at a putative transformant after 10 days (D), and closer view of the transformant (E). (C, F) Putative transformants grown on soil after 1 month.

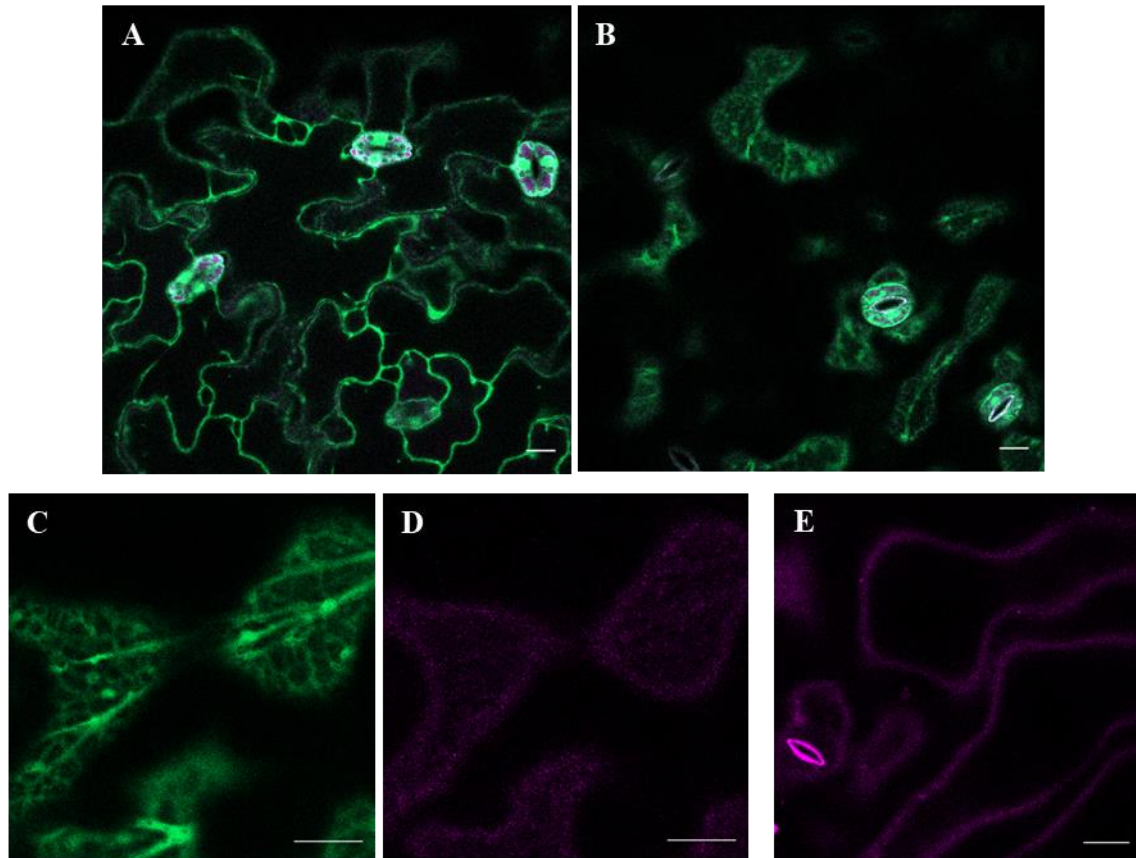


Figure 6.2. Confocal microscopy images of T1 plants.

(A, B) Low magnification images of T1 plant at flowering stage following transformation with RFP-polypeptide. (C, D) Higher magnification image of T1 plant transformed with RFP-polypeptide showing expression of clover in the cytosol, and no expression of RFP-pmoC. (E) Image of T1 plant at flowering stage transformed with RFP-CAB showing no expression of RFP-pmoC. Scale bar =10 μ m.

T2 plants of RFP-polypeptide were generated to obtain homozygous lines and to image the seedlings at a younger stage, to assess whether the silencing of RFP-pmoC was already occurring at an early growth stage. Germination of the T2 seeds in selective media produced a germination proportion of 18/48 (0.375), which is lower than the 0.75 proportion expected in the case of mendelian segregation (Figure 6.3 A). However, T2 seeds grown without antibiotics had a germination proportion lower than 100% with 29/48 seedlings growing on MS media (0.604) (Figure 6.3 B). When considering the germination rates with and without antibiotics, the former value approximates the mendelian proportion more closely. WT *Arabidopsis* seeds did not germinate in the selective media, confirming the efficacy of the antibiotics used for

selection (Figure 6.3 C). WT seeds grown in MS media without antibiotics also showed a germination rate lower than 100% (Figure 6.3 D). This could be due to the WT seed aliquot being kept for a long period of time, rendering some seeds unviable, or it could be due to some seeds losing their germination capacity during their sterilization with 70% ethanol prior to plating. At 10 days old the RFP-polypeptide seedlings grown with antibiotics did not show apparent phenotypic defects compared with WT seedlings grown without antibiotics, if anything they appear to be slightly larger than the WT seedlings. Quantification of the area of individual seedlings yielded an average area of $15 \pm 4 \text{ mm}^2$ (mean $\pm$ SD) for RFP-polypeptide in the selective media, $16 \pm 6 \text{ mm}^2$ (mean $\pm$ SD) for RFP-polypeptide in MS media without antibiotics, and $11 \pm 4 \text{ mm}^2$ (mean $\pm$ SD) for WT seedlings without antibiotics (Figure 6.3 E). Seedlings with an area below 5 mm^2 were excluded from the quantification.

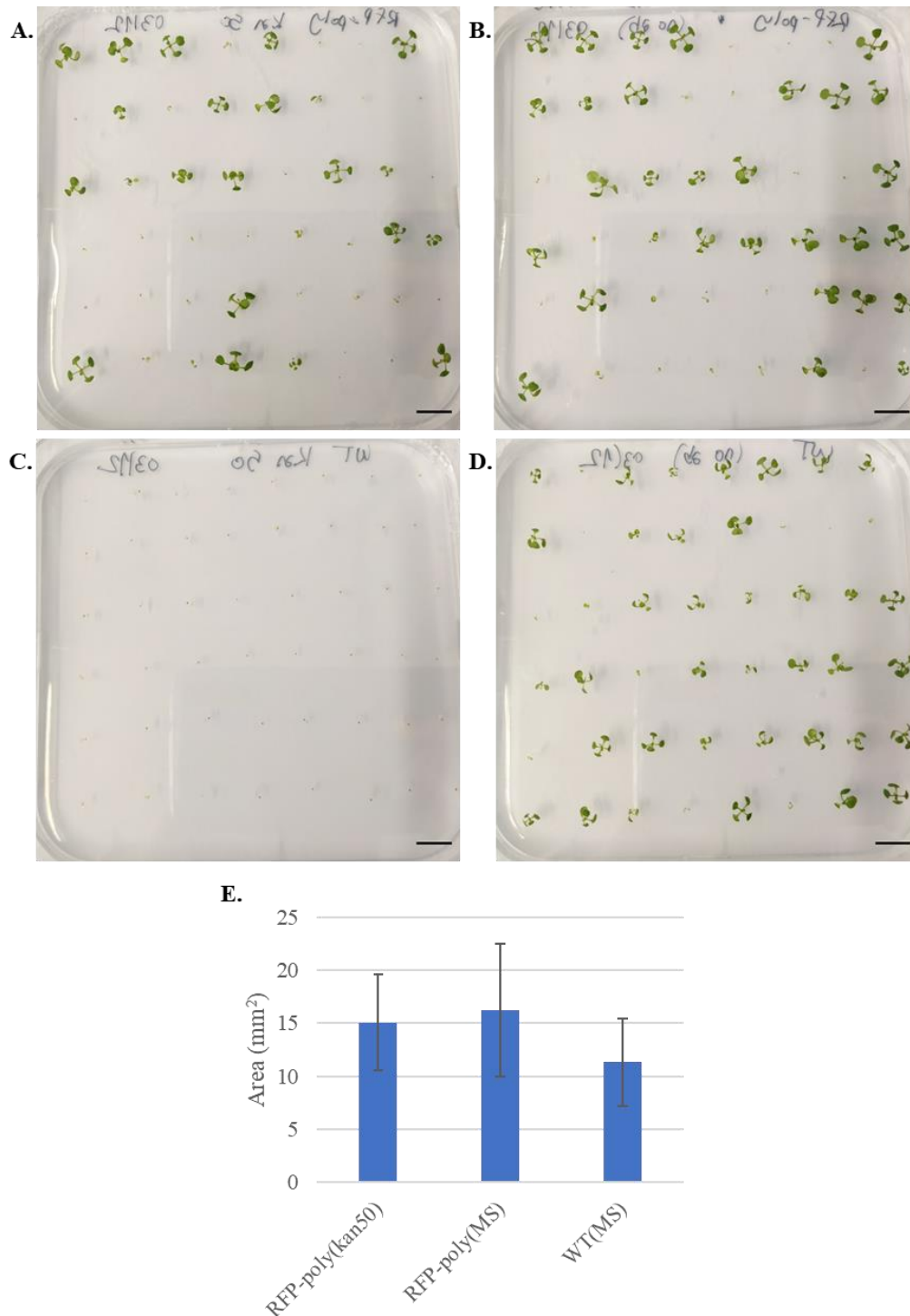


Figure 6.3. Germination of T2 seeds on selective media.

10 day old Arabidopsis seedlings of (A) T2 RFP-polypeptide on MS media with kanamycin (50 mgL⁻¹), (B) T2 RFP-polypeptide on MS media without antibiotics, (C) WT Arabidopsis on MS media with kanamycin (50 mgL⁻¹), and (D) WT Arabidopsis on MS media without antibiotics. Scale bar = 10 mm. (E) Quantification of seedling size of T2 RFP-polypeptide with kanamycin (n=18), T2 RFP-polypeptide without antibiotics (n=26) and WT without antibiotics (n=26). Showing the mean and standard deviation.

5 day old seedlings were imaged and showed strong expression of clover in the cytosol, in cotyledons (Figure 6.4 A, B), root cells (Figure 6.4 C), and hypocotyl (Figure 6.4 D). Additionally, a faint signal in the red channel labelling the ER could be seen, indicating expression of RFP-pmoC. This signal was very close to background levels, and upon adjusting the brightness level, the presence of the ER structure in the labelling becomes clear. This indicated that at a very young stage in development Arabidopsis plants had not yet fully silenced the pMMO part of the polypeptide, and expression of RFP-pmoC could be detected, albeit at very low levels.

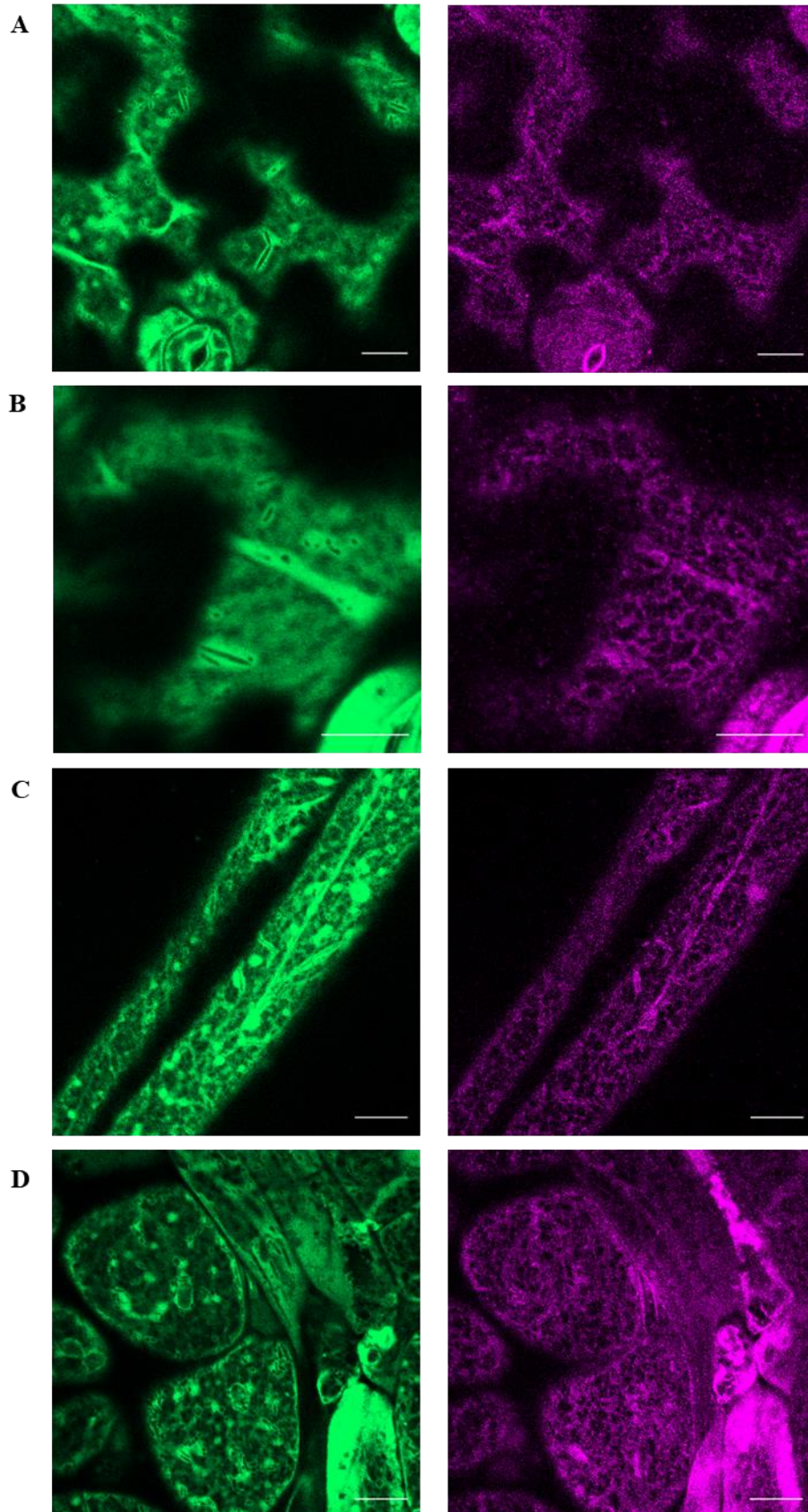


Figure 6.4. Expression of RFP-polypeptide in 5-day old Arabidopsis seedlings.

Confocal images of 5-day old Arabidopsis T2 seedlings expressing RFP-polypeptide showing the green channel (clover) and red channel (RFP-pmoC), in (A and B) cotyledons, (C) roots and (D) hypocotyl. Scale bar =10 μ m.

6.2.2 Analysis of the impact of protein additions in the polypeptide

Given stable Arabidopsis lines only showed low expression, further steps were taken to maximize the likelihood of successful reconstitution of pMMO in transient assays. The polypeptide used in this study included short protein additions to the subunits derived from the 2A self-cleaving peptides. Analysis of the quaternary structure of pMMO was carried out to evaluate whether these additions could potentially interfere with the complex assembly, and consequently hinder enzymatic function. Modelling approaches were attempted to predict the structure of the subunits with the added residues (20, 21 and 22 residues added to the C-terminus of pmoC, pmoA and spBtrunc respectively). One approach used SWISS-MODEL, which carries out homology modelling based on templates from known structures found in the PDB database. However, this approach only produced a structure for the regions of the sequence which aligned sufficiently with the template, and the protein additions were not modelled. Other bioinformatic tools used were I-TASSER and Phyre², which combine template based modelling with *ab initio* folding, for regions without a template. These software packages produced full models of the subunits with the additions, however, the confidence score for the additional regions was extremely low. *Ab initio* modelling is considered to be unreliable for protein structure prediction in comparison to template based models (Lee *et al.*, 2017). This is especially true for highly flexible regions which can adopt multiple conformations, such as is the case for the self-cleaving peptides, which are short and have a high content of glycine, serine and alanine residues. Therefore, modelling approaches were not considered appropriate to provide a confident result for the folding of the subunits with the protein additions. Instead, the quaternary structure of pMMO was examined using PyMOL to evaluate whether the additions would be located at sterically restricted regions which could interfere with the native folding.

The pMMO crystal structure from *Methylosinus trichosporium* OB3b (PDB 3CHX), used in this study, shows the C-terminus of pmoC facing the cytosolic part of the structure, and located on the outer perimeter of the nonamer (Figure 6.5 A). However, this structure contains 80 unmodelled residues at the C-terminus. Therefore, the pMMO structure of *Methylocystis* sp. strain M (PDB 3RFR) was analyzed, as it contains the highest sequence similarity with that of *Methylosinus trichosporium* OB3b out of the other PDB structures available. The final residues of the pmoC subunit are modelled in this structure and show the C-terminus to be located facing the outside of the nonamer, therefore suggesting that protein additions on this location might not compromise interactions with the other subunits (Figure 6.5 B).

The C-terminus of pmoA is found to be located towards the interior of the nonamer structure in 3CHX (Figure 6.5 C). However, topology and protein-protein interaction data suggested that the C-terminus of pmoA is folded differently and faces the cytosol (see Chapter 3). Hence, the 3RFR structure was analyzed. The C-terminus of pmoA in 3RFR can be seen in the cytosolic face of the enzyme, although in the inner perimeter, and in very close proximity to pmoC, and to pmoB of the adjacent trimer (Figure 6.5 D). This suggested that protein additions on this position might interfere with the ability of pmoA to interact properly in the complex.

Finally, the C-terminus of the pmoB subunit can be seen in the luminal section of the structure facing outwards from the complex, suggesting that additions to this terminus would not have a big impact upon protein-protein interactions (Figure 6.5 E). On the other hand, the N-terminus of pmoB is located on a Copper site, at the interface between the pmoB beta barrels and pmoC, with the N-terminal histidine being one of the three coordinating residues of the copper atom (Figure 6.5 F). Additions to this terminus could potentially disrupt the coordination with copper. Even though no amino acids were added to the N-terminus by self-cleaving peptides, three amino acids (YRT) were added by the plant signal peptide used to create spBtrunc, due to the location of the cleavage site, therefore, this was considered for further bioengineering strategies.

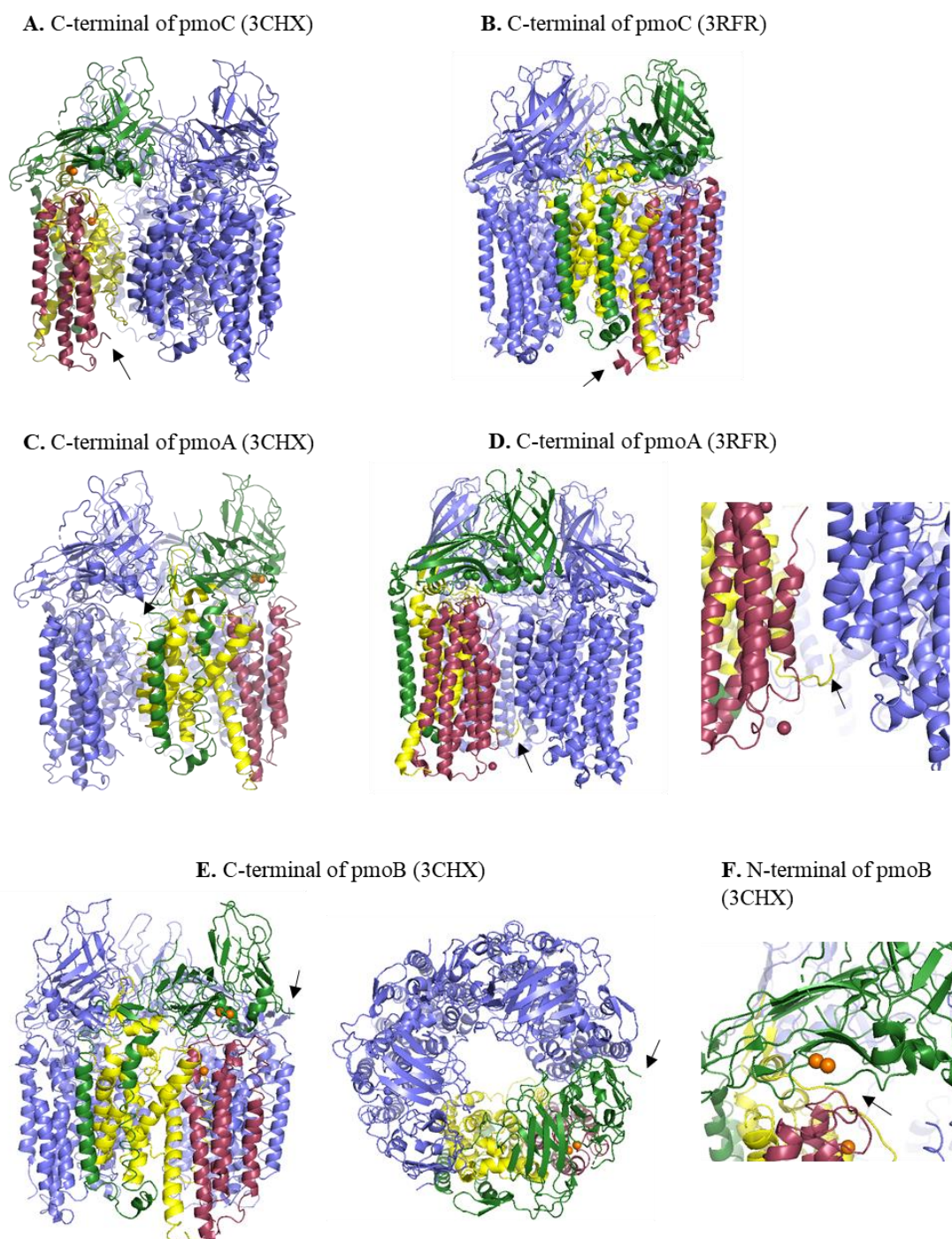


Figure 6.5. Position of the N- and C- termini of pMMO subunits in the crystal structure.

The structures of the pMMO nonamer from *Methylosinus trichosporium* OB3b (PDB ID 3CHX) and *Methylocystis* sp. strain M (PDB ID 3RFR) are shown. For one trimer pmoC is shown in red, pmoA in yellow, and pmoB in green, the remaining two trimers are shown in purple. (A) The C-terminus of pmoC in structure 3CHX is shown with a black arrow. (B) The C-terminus of pmoC in structure 3RFR is shown, which has additional modelled residues. (C) The C-terminus of pmoA in structure 3CHX is shown, pointing towards the inside of the nonamer in the lumen. (D) The C-terminus of pmoA in structure 3RFR is shown pointing towards the inside of the nonamer in the cytoplasm, in accordance with the topology found in

this study. (E) The C-terminus of pmoB from structure 3CHX is shown pointing towards the outside of the nanomer in both a side view and a top view of the structure. (F) The N-terminus of pmoB from 3CHX faces the inside of the structure and participates in the coordination of the copper site.

6.2.3 Design and transient expression of a modified polypeptide for increased functionality

The analysis of the location of the protein termini within the pMMO structure suggested that residue additions to the N-terminus of pmoB and to the C-terminus of pmoA may be unfavourable to reconstitute native interactions and binding to copper. In order to avoid the addition of three residues to the N-terminus of spBtrunc, derived from the plant signal peptide which was used to target pmoB to the ER, a construct was cloned using a different plant signal peptide, from tobacco pathogenesis-related protein 1a (PR1a), which has been used in other studies to target recombinant proteins to the secretory pathway. The cleavage site of this signal peptide is located between its last residue and the first residue of the protein (Figure 6.6 A), leaving no additional residues at the N-terminus of the protein. The new construct, termed PR1Btrunc was cloned with a C-terminal GFP and transiently expressed in tobacco. Expression could be seen in the endoplasmic reticulum, showing correct targeting by the new signal peptide, and expression levels were very low as was the case for spBtrunc-GFP (Figure 6.6 B).

In combination with the use of PR1Btrunc, the polypeptide sequence was further modified to exclude additions to the C-terminus of pmoA, by adding an intein domain upstream of the P2A peptide (Figure 6.6 C). The intein domain possesses hyper-N-terminal autocleavage capacity and if fused upstream of a 2A peptide constitutes a self-excising fragment which leaves no added residues to the proteins (Zhang *et al.*, 2017).

A final modification was made with the aim of improving expression levels of the polypeptide. A small protein domain, termed the “M domain”, which harbours several N-glycosylation sites, was reported to greatly improve the expression levels of recombinant proteins targeted to the ER (Kang *et al.*, 2018). Even though it was seen that substantial polypeptide downregulation occurs through PTGS mechanisms, as its signal increases by co-expression with p19 (see

Chapter 4), other mechanisms such as protein degradation may also be contributing to the low expression levels. Therefore, the M domain was fused to the C-terminus of PR1Btrunc, in the hope that the N-glycosylation moieties would improve the protein stability or prevent degradation. The new polypeptide containing the three aforementioned modifications was termed “polyM” (Figure 6.6 C).

Figure 6.6. Design of the modified polypeptide polyM.

The polyM sequence was codon optimised and synthesized, and cloned into plant expression vectors containing either an N-terminal RFP or a C-terminal GFP fluorophore, yielding RFP-polyM and polyM-GFP, respectively (Figure 6.7 A). Transient expression in tobacco showed very low expression levels of both constructs (Figure 6.7 B, D), as was the case with the original

polypeptide, indicating that the M domain did not confer any detectable increase in protein quantities. Co-expression with p19 produced a marked increase in expression (Figure 6.7 C, E), also in a similar manner to the original polypeptide. Localization was seen in the ER for both RFP-polyM and polyM-GFP, showing that the modifications to the polypeptide sequence did not alter the ER targeting of the subunits.

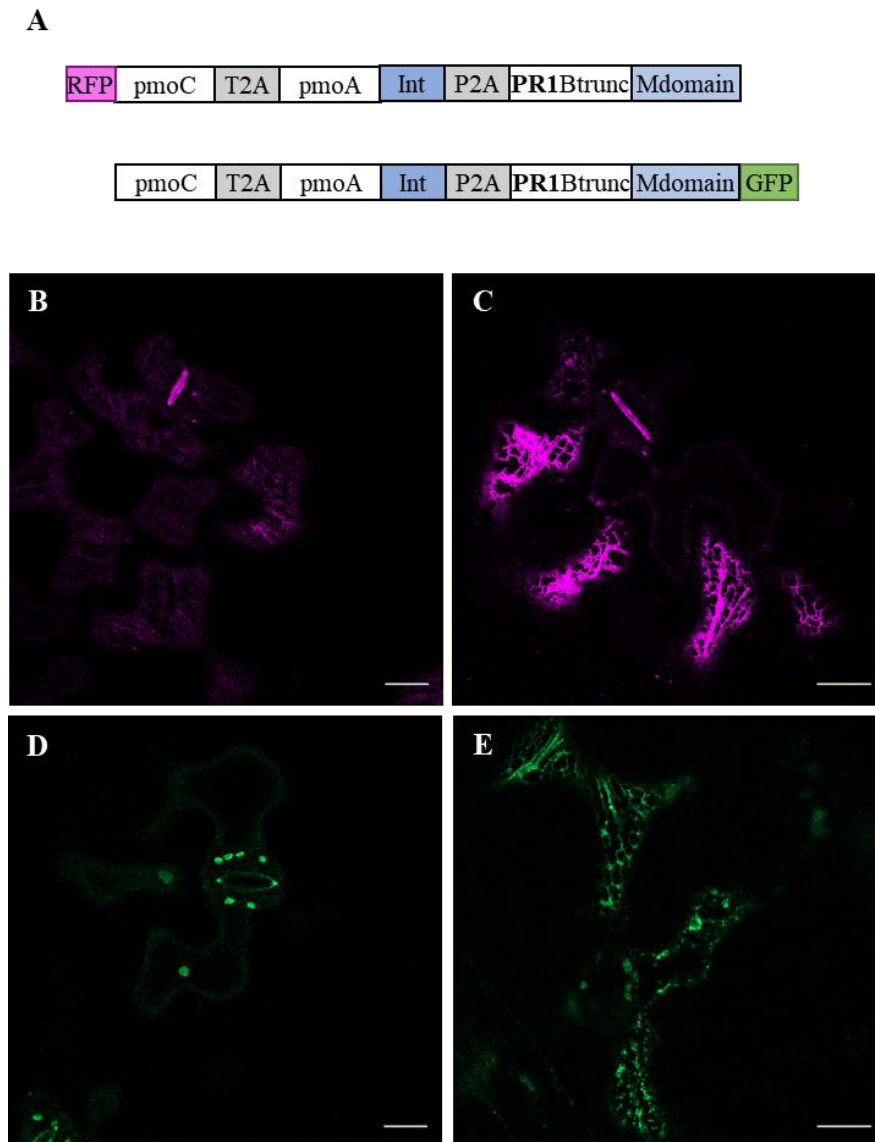


Figure 6.7. Transient expression of RFP-polyM and polyM-GFP.

(A) Schematic diagram of polyM with fluorescent protein fusions. (B) Expression of RFP-polyM in tobacco epidermal cells at 2 days after infiltration, (C) co-expression of RFP-polyM with p19, (D) transient expression of polyM-GFP at 2 days after infiltration, (E) co-expression of polyM-GFP with p19. Scale bar = 20 μ m

6.2.4 Design of propylene epoxidation assays and sample preparation

Propylene epoxidation assays were set up to test activity of pMMO in plant samples *in vitro*, where activity would be detected by the conversion of propylene gas to propylene oxide. Given the very low abundance of the enzyme in stable *Arabidopsis* seedlings, these samples were not included for a first trial of enzyme activity. Instead, transient expression of the polypeptide and of polyM with p19 yielded the most favourable strategy to obtain high yields of the polypeptide. Therefore, RFP-poly and RFP-polyM (containing no additions to the C-terminus of pmoA and N-terminus of Btrunc) transiently expressed in *Nicotiana benthamiana* with p19 were used as samples for the first trial of activity.

Plants were infiltrated and lysates were made two days later by grinding the infiltrated tissue with liquid nitrogen and extracting proteins using a non-denaturing lysis buffer (Figure 6.8). The extraction was also carried out with the lysis buffer containing CuSO₄, to approximate the method used in the assay for methanotroph cells, where high levels of copper are required for expression and stability of pMMO. Following centrifugation at low rpm, the supernatant was snap frozen and stored at -80 °C for future use in assays. Assays in methanotrophs require the addition of duroquinol as a reducing agent to test the activity of pMMO *in vitro*. It was not known whether any native plant molecules would act as reducing agents for pMMO in the plant ER. Therefore, the assay plan included samples both with and without duroquinol (Table 6.1). The enzymatic assays were to be carried out in collaboration with Professor Tom Smith at Sheffield Hallam University, and due to COVID-19 restrictions and technical issues, the assays have not been carried out to date. However, the experimental design and procedure for the assays was established (Figure 6.8 and Table 6.1). The sample and reagent preparation was also carried out.

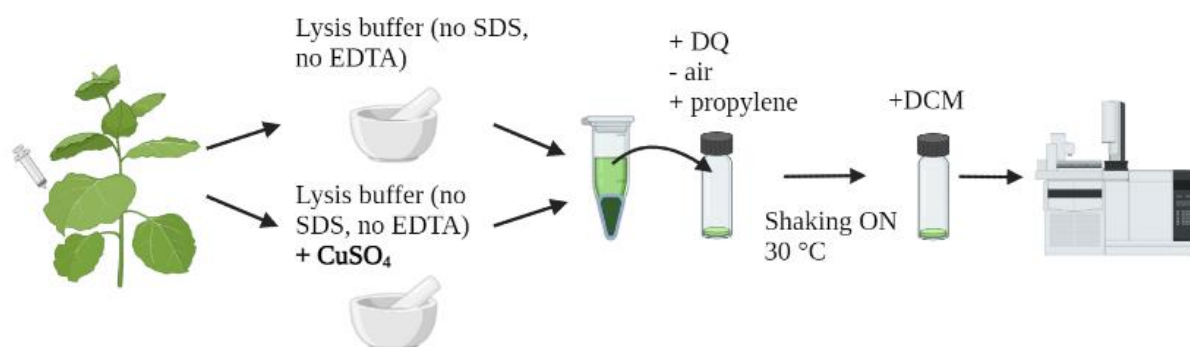


Figure 6.8. Simplified diagram of the proposed process for propylene epoxidation assays from plant extracts.

Plants are infiltrated and after 2 days leaf pieces are ground with liquid nitrogen to a fine powder. Lysis buffer is added containing no SDS or EDTA (2 common components of plant lysis buffers) to avoid protein denaturation and chelation of copper ions, respectively, with and without CuSO_4 . The lysates are centrifuged at low rpm to remove large tissue fragments, and an aliquot of the supernatant is placed in a glass vial, to which duroquinol (DQ) is added. The glass vials are sealed and air is removed with a syringe and replaced with propylene gas. The vials are shaken overnight at 30°C, and are then cooled in a fridge to avoid any evaporation of the propylene oxide which is a volatile liquid. Dichloromethane (DCM) is used as a solvent to extract the propylene oxide. A small volume from the clear, organic phase is inserted into the gas chromatography equipment for analysis.

Table 6.1. Experimental design of enzyme assays.

Extracts from wild type (WT) *Nicotiana benthamiana*, or *Nicotiana benthamiana* tissue expressing RFP-poly (OD_{600} 0.1) + p19 (OD_{600} 0.1), or RFP-polyM (OD_{600} 0.1) + p19 (OD_{600} 0.1) are used, and *Methylosinus trichosporium* OB3b are included as a positive control. Samples are tested with and without the addition of 10 mM duroquinol (DQ), and with and without the inclusion of 40 μM CuSO_4 during lysis (the duroquinol and copper concentrations were taken from the values used in (Kitmitto *et al.*, 2005)).

1. WT		
2. RFP-poly + p19		
3. RFP-polyM + p19		
4. WT	+ 10 mM DQ	
5. RFP-poly + p19	+ 10 mM DQ	
6. RFP-polyM + p19	+ 10 mM DQ	
7. WT	+ 10 mM DQ	+ 40 μM CuSO_4
8. RFP-poly + p19	+ 10 mM DQ	+ 40 μM CuSO_4
9. RFP-polyM + p19	+ 10 mM DQ	+ 40 μM CuSO_4
11. <i>Methylosinus trichosporium</i> OB3b	+ 10 mM DQ	+ 40 μM CuSO_4

6.2.5 Transgenic rice containing the pMMO polypeptide

The principal application for expressing pMMO in plants was to create transgenic rice which could mitigate methane emissions from paddy fields. Even though the results from activity assays in tobacco were not yet obtained, the transformation of rice plants containing the polypeptide was carried out in parallel given the long time frames required to produce transgenic rice. Rice plants (cv. Nipponbare) were transformed with the polypeptide sequence to evaluate the enzyme localization and functionality in the target crop species. The polypeptide sequence (the sequence in Appendix 1 with an added N-terminal ribosome binding site: ccacc) was cloned into a pDONR vector and sent to the National Institute of Agricultural Botany (NIAB), where it was cloned into a plant destination vector. Rice transformation was performed at NIAB and yielded 39 transformants and 2 wild-type controls. Transformation was confirmed at NIAB using qPCR, determining the copy number in each line. Following delivery of the seedlings in Jiffy Pellets the seedlings were transplanted into clay granules and grown in tropical greenhouse conditions. The lines presented highly varied growth phenotypes, with some plants reaching a large size and tiller number at two months after transplant, whereas other plants presented intermediate growth levels, and others stunted growth or chlorosis (Figure 6.9). Unexpectedly, the two control plants had highly stunted growth, suggesting that this phenotype could be due to stressful conditions during the transport and transplant or by the presence of pathogens or lack of nutrients. In order to characterize the phenotype of pMMO plants, a larger set of control plants would be required to control for stochastic events during growth, as well as avoiding transportation and transplanting. This could be carried out in T1 plants. Nevertheless, by observation of the T0 generation, it could be concluded that rice plants containing the pMMO polypeptide were able to grow and produce seeds (Figure 6.9 G, H).



Figure 6.9. Images of rice plants at two months after transplant.

(A-F) Photos showing the growth of different transformed lines and a control line (C1). Images were taken by John Baker (Department of Plant Sciences, University of Oxford). Ruler = 10 cm. (G, H) A transformed line (11B) at maturity in the grain filling stage, showing the presence of several panicles (G) and a closer view of the grains (H).

RT-PCR was carried out on tissue material from the leaf blade obtained from plants at two months after transplant, to assess transcription of the polypeptide. An untransformed control line (C1) was included, alongside a subset of 11 transformed lines, and a control for DNA contamination without reverse transcriptase (no RT control) for line 3B (Figure 6.10). Amplification with actin primers was used as a control for the presence of RNA in the extraction. Amplification with pmoC, pmoA, and pmoB specific primers showed a faint band in most lines (excluding the control), and a stronger band in all three subunits in the 3B line (Figure 6.10 A-C). The control without reverse transcriptase (no-RT control) of the 3B line did not produce a detectable band in any of the subunits. This showed that there was no DNA contamination and therefore that line 3B was indeed transcribing the polypeptide genes. Moreover, line 3B showed a large size and number of tillers at the time of sampling (Figure 6.10 B), indicating that the transcription of pMMO genes was not detrimental to the growth of rice plants. Further no-RT controls of the remaining lines would be required to confirm transcription of the genes in those plants.

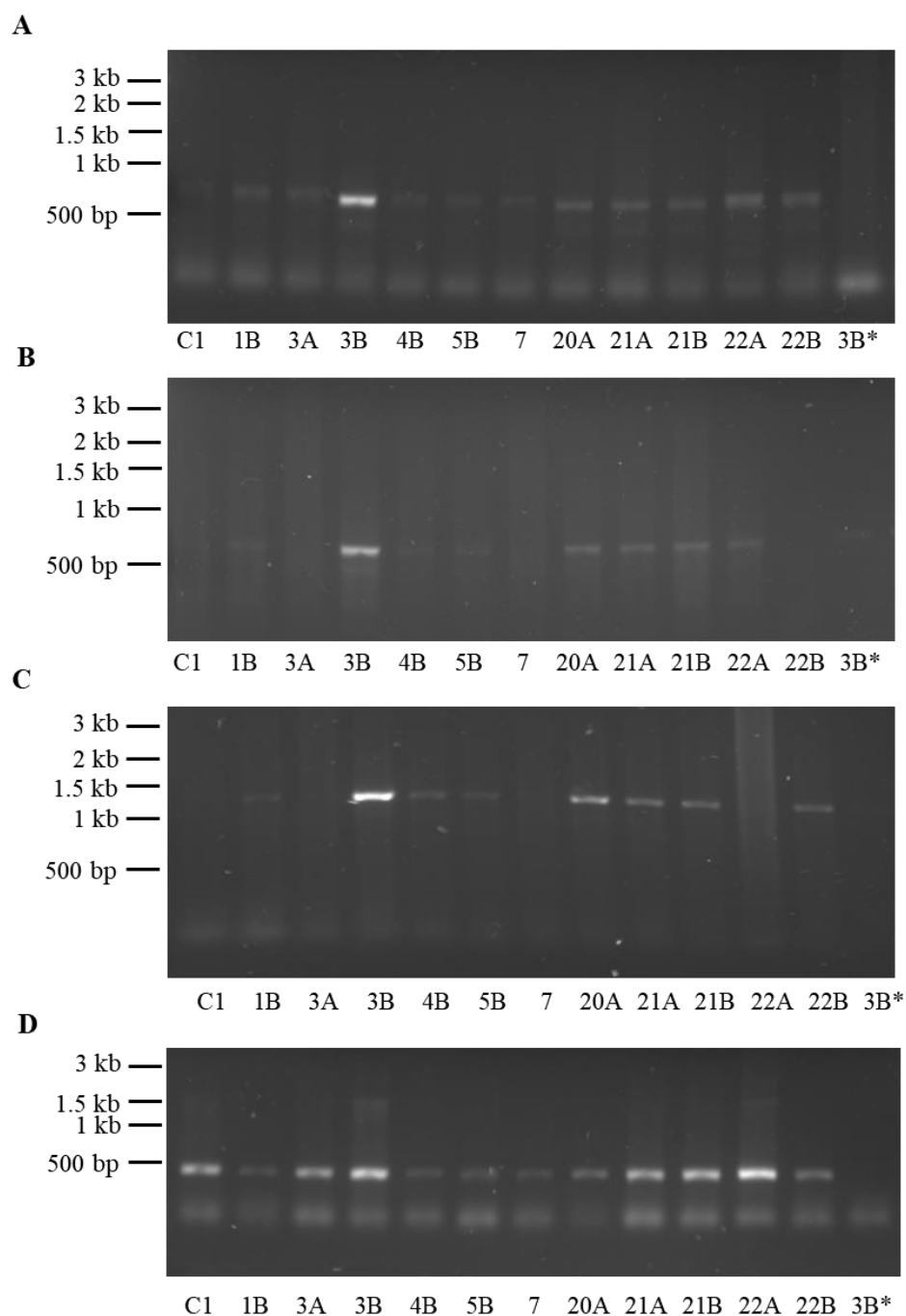


Figure 6.10. Transcription of pMMO polypeptide genes in rice.

RT-PCR from leaf tissue of rice plants 2 months after transplant, in an untransformed plant (C1, control 1) and 11 transgenic lines (1B, 3A, 3B, 4B, 5B, 7, 20A, 21A, 21B, 22A, 22B). 3B* indicates no-RT control of line 3B. Amplification of cDNA samples was carried out with primers for pmoC (expected size 757 bp) (A), pmoA (expected size 748 bp) (B), pmoB (expected size 1247 bp) (C) and actin (expected size 435 bp) (D).

Some of the lines were imaged using confocal microscopy to evaluate translation of the polypeptide containing the terminal clover fluorophore. However, confocal imaging of the rice leaf blade is difficult because of the high level of light scattering in the waxy cuticle (Luginbuehl *et al.*, 2020). Imaging was also complicated by the parallel and compacted veins present throughout the leaf blade, which rendered regions of the ROI to be alternatively in and out of focus, as well as by the diversity of cell types present. The cell types were identified based on labelled images present in (Luo *et al.*, 2012). Some polypeptide transgenic lines such as 1A and 5A, showed large growth and a deep green coloration of the leaf blade at the time of imaging, whereas other lines such as 4A and the control line showed a much lighter coloration (Figure 6.11 A). Confocal imaging of line 1A showed a strong signal in the GFP channel in the regions adjacent to a row of short cells (with gaps in the green signal appearing in stomata) (Figure 6.11 B). This suggested translation of the clover protein.

The transgenic line 4A also showed some signal in the GFP channel in the regions adjacent to rows of short cells (Figure 6.12 A). Some signal in the GFP channel was also observed in rows of long cells, and in puncta deriving from long cells, although it is not clear if this signal derives from autofluorescence (Figure 6.12 B). The control line did not show any green signal adjacent to what appears to be similar rows of short cells as were seen in the transgenic line (Figure 6.12 C and D). However, autofluorescence could clearly be seen in the control line along different rows of short cells, and in puncta of long cells (Figure 6.12 E), similar to the puncta seen in the transgenic line (Figure 6.12 B). Therefore, the transgenic lines are potentially translating clover from the polypeptide, but it is unclear whether this signal is markedly distinct from autofluorescence.

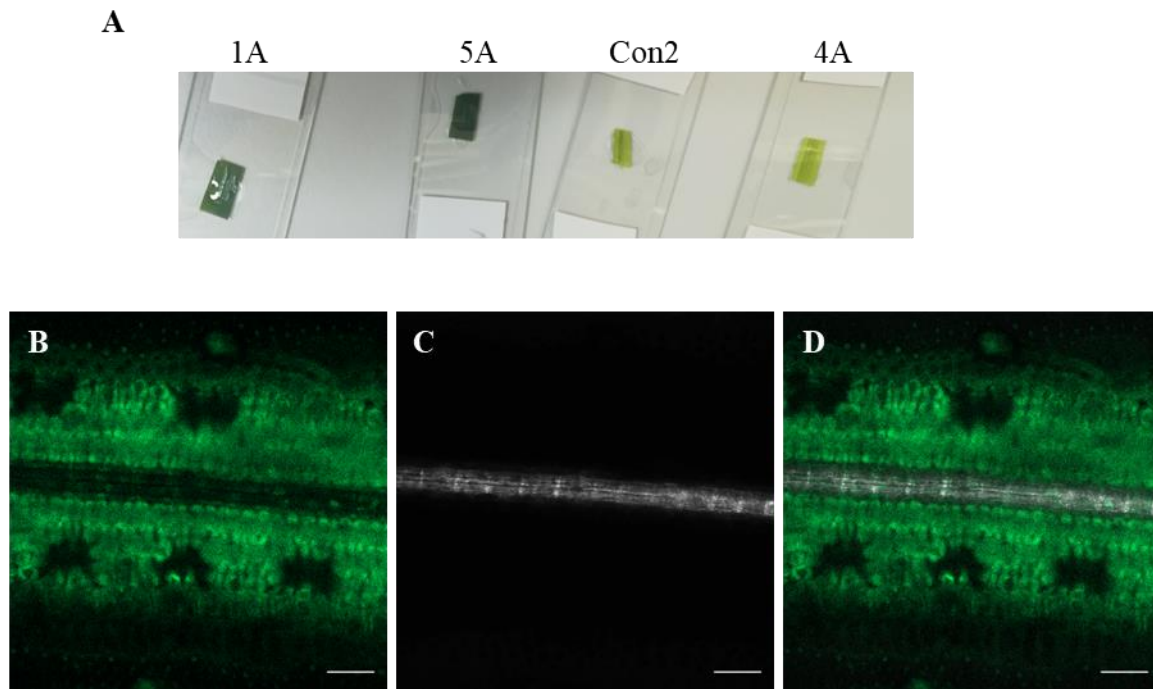


Figure 6.11. Confocal imaging of transgenic line 1A containing the pMMO polypeptide.

(A) Photos of sections from the leaf blade of transgenic rice plants 1A, 5A, 4A, and an untransformed line (Control 2), showing different chlorophyll content. Confocal imaging of line 1A showing the GFP channel (B), transmitted light (C), and merged channels (D), with signal in the GFP channel signal adjacent to a row of short cells which could derive from the translation of clover. Scale bar = 20 μm .

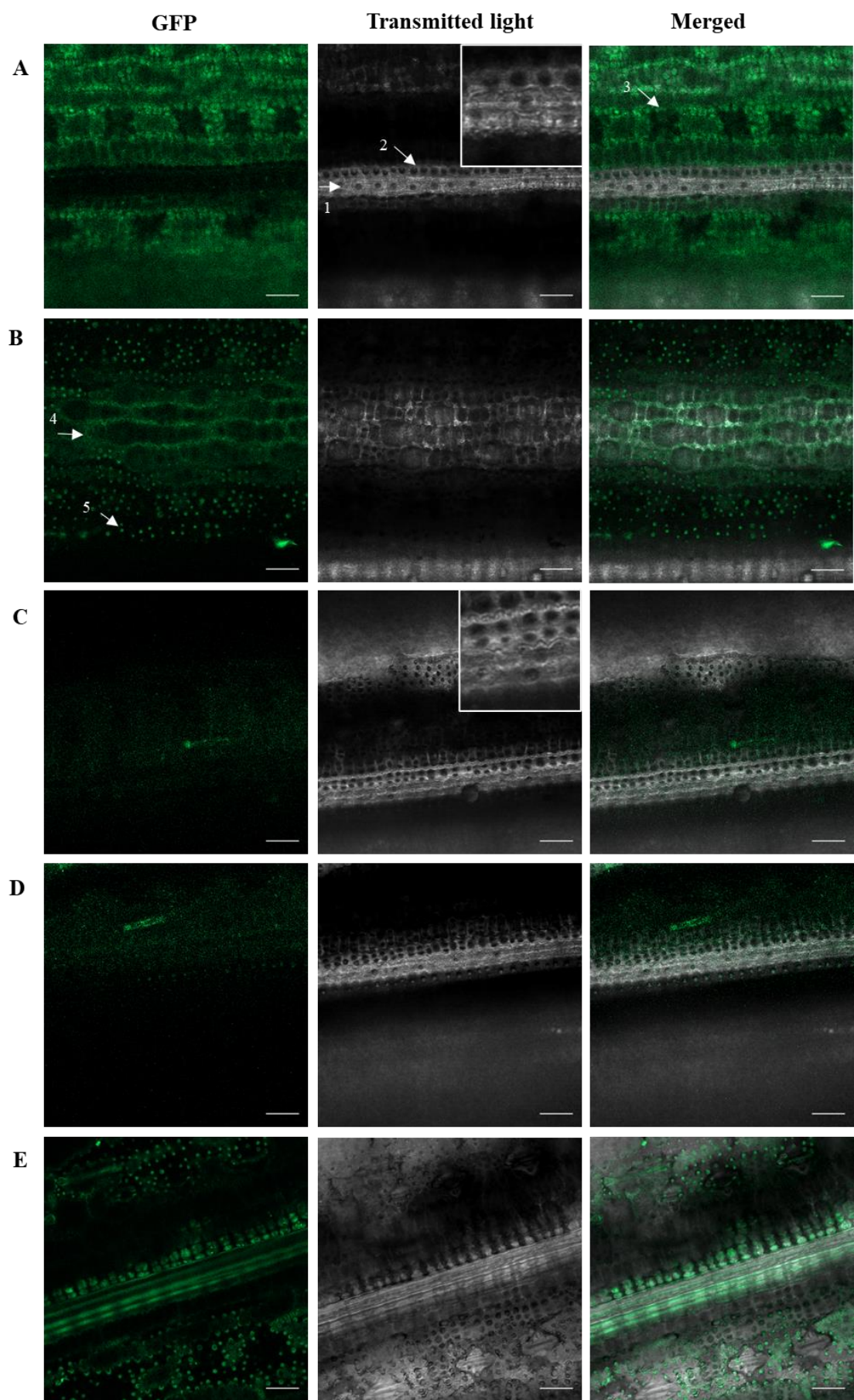


Figure 6.12. Confocal imaging of rice leaf blades.

(A) Line 4A containing the pMMO polypeptide showing possible expression of clover in the GFP channel adjacent to a row of short cells, inset showing a closer view of the short cells. (B) Line 4A showing signal or potential autofluorescence in the GFP channel. (C, D) Rows of short cells in a control line with no signal in the GFP channel, inset showing a closer view of the short cells. (E) Short cells and long cells bordering stomata in the control line, showing autofluorescence in the GFP channel. 1= Row of short cells, 2= row of long cells flanking short cells, 3= stomata, 4 =ordinary long cells, 5= long cells in stomata files. Scale bar = 20 μ m.

6.3 Discussion

6.3.1 Creation of stable *Arabidopsis* plants containing pMMO

Transgenic *Arabidopsis* plants containing the RFP-polypeptide showed expression of clover, present at the C-terminus of the polypeptide, but no detectable expression of RFP-pmoC, present at the N-terminus of the polypeptide, at the adult stage. Silencing in *Arabidopsis* T-DNA transformants is a common occurrence, with marked differences in expression levels often found between different transformant lines. This is due to the randomness of the insertion process and copy number, which can trigger PTGS or TGS, therefore the screening of multiple lines is generally carried out (Schubert *et al.*, 2004; Butaye *et al.*, 2004). In this case, transient expression of the RFP-polypeptide in tobacco already showed strong downregulation via PTGS, where only clover could be seen in most cells, and faint traces of RFP-pmoC in some cells (see Chapter 4). Therefore, the silencing observed in the RFP-polypeptide *Arabidopsis* line is likely independent of position effects arising during integration, and is likely to be observed in any additional transgenic lines created. Interestingly, at a very early growth stage expression of RFP-pmoC could be seen in the ER in the different organs of the seedling (Figure 6.4). Additionally, the expression of RFP-polypeptide did not produce apparent defects in germination rate or growth.

6.3.2 Testing enzymatic activity of pMMO in plants

The expression of pMMO in stable *Arabidopsis* plants resulted in very low protein levels, where expression of RFP-pmoC could not be detected at the adult stage, and could be detected at trace levels at the seedling stage. Therefore, it was chosen to carry out a first trial of enzymatic activity using transiently expressed RFP-polypeptide with p19 in *Nicotiana benthamiana*, as this proved to be the strategy which induced higher expression levels of the construct. Additionally, transiently expressed RFP-polyM with p19 was included in the assay design to increase the likelihood of achieving enzymatic activity by removing certain amino

acid additions deriving from the self-cleaving peptides. Propylene epoxidation assays using plant lysates were chosen as a method that could provide sufficient sensitivity to detect enzymatic activity even if present at low levels. Testing enzyme activity of pMMO has proven difficult in several laboratories, with groups supplementing the growth medium of methanotrophs with copper ions to stimulate the formation of ICMs, the expression of pMMO and stability of pMMO in the membranes (Nguyen *et al.*, 1994; Choi *et al.*, 2003; Kitmitto *et al.*, 2005; Smith *et al.*, 2011). Some studies in methanotrophs also add supplemental copper ions to the membrane fraction or to the assay medium to achieve higher levels of enzymatic activities *in vitro* (Nguyen *et al.*, 1994; Takeguchi *et al.*, 1998; Kitmitto *et al.*, 2005; Hakemian *et al.*, 2008). Copper is an essential micronutrient in plants, and is necessary for cellular processes such as the electron transport chains in photosynthesis and in cellular respiration, ethylene sensing, and responses to oxidative stress, amongst others (Yruela, 2009). However, high levels of copper are toxic in plants (Yruela, 2009), so it is expected that only trace levels of copper will be present in the plant lysates, and supplementation of copper could be necessary to stabilize the pMMO complex. CuSO₄ was thus added to the lysis buffer to potentially stabilize pMMO in the ER membrane and during the assay.

Another additive to the assay was duroquinol to act as a reducing agent. pMMO can use NADH when assays are carried out in whole-cells or membrane fractions, but not with solubilized or purified pMMO. On the other hand, duroquinol, a synthetic quinol, is effective at eliciting activity of pMMO in whole-cells, membrane fractions and solubilized pMMO (Shiemke *et al.*, 1995). It is thought that the enzyme in methanotrophs thus uses a quinol as a reducing agent, such as ubiquinol, which is produced by NDH-2 in the methanotroph respiratory chain (Sirajuddin and Rosenzweig, 2015), however, the physiological reductant of pMMO has not been confirmed to date. Besides from duroquinol, decyl-plastoquinol also served as an electron donor for pMMO assays in whole cells and in solubilized pMMO, whereas ubiquinol did not (Table 6.2). Plastoquinol and ubiquinol are present in plant cells in the chloroplast thylakoid membrane and the mitochondrial inner membrane, respectively (Liu and Lu, 2016). However,

pMMO is present in the ER membrane in the plant lysates, therefore it is unlikely that it would be reduced by any native quinols, or that NADH or NADPH could act as indirect reductants. The addition of duroquinol thus seems necessary in this case to elicit pMMO activity in the lysates. Assaying the samples with and without duroquinol would provide confirmation of this requirement.

Table 6.2. The effect of different reductants on enzymatic activity of pMMO.

Specific activity (nmol min⁻¹ mg⁻¹) of different samples containing pMMO using various reducing agents, measured via propylene epoxidation assay. Cropped from (Shiemke *et al.*, 1995). The assays with higher activity values are highlighted in green.

Reductant	Whole cells	Membranes	Soluble fraction
NADH	24.2	6.2	<0.1
Decyl-plastoquinol	4.44	0.38	2.70
Duroquinol	2.90	1.56	1.46
Decyl-ubiquinol	0.41	0.18	0.36
Coenzyme Q ₀	1.22	0.08	0.18

6.3.3 The generation of rice containing pMMO

Rice transformation performed at NIAB with the polypeptide sequence produced several transformant lines. These lines exhibited very variable growth phenotypes, however, due to stunted growth in the control lines, it is likely that many plants suffered during transplant, with the change in growth conditions, or with the advent of pests during growth, inducing the different phenotypes. A study of T1 plants with an equal number of control plants, grown without transport and appropriate conditions would allow a more accurate phenotypic characterization of the transformants. Nevertheless, many T0 transgenic lines reached a large size and tiller number during growth, and some plants produced seeds. Short day conditions were necessary for the plants to produce seeds. One line (3B) clearly showed transcription of the pMMO genes at the adult stage (Figure 6.10). Therefore, the results indicate that rice plants transcribing the pMMO genes are able to grow, and importantly produce grains. In terms of evaluating translation of the polypeptide, confocal imaging was carried out to detect the C-terminal clover, however, these results were not conclusive (Figure 6.12). Confocal imaging of

rice plants is difficult due to the waxy cuticle in the leaf blade producing high levels of light scattering, and difficulty in distinguishing autofluorescence from actual signal (Luginbuehl *et al.*, 2020). Recently protocols such as the ClearSee protocol or scraping the leaf surface, together with appropriate reporter fluorophores targeted to the nucleus, have been optimized for clearer imaging of rice leaf blades (Luginbuehl *et al.*, 2020). These techniques could be applied to determining translation of the polypeptide in rice plants. However, due to the results in tobacco where the clover fluorophore was expressed but not the pmoC subunit, Western blots would provide a more robust method for determining translation of the subunits.

7 General discussion

7.1 Expression of pmoC, pmoA and pmoB in the plant ER

Recombinant expression of the pMMO subunits is a largely unexplored subject in the literature due to limited successes at heterologous expression. One study has expressed the whole pMMO complex in *Rhodococcus erythropolis* (Gou *et al.*, 2006), a Gram-positive bacteria studied for biotechnology purposes due to its ability to bioremediate diverse environmental pollutants. However, the activity levels of the recombinant strain were extremely low and no experiments were made regarding the correct membrane localization of the complex or the interactions between the subunits (Gou *et al.*, 2006). Expression of pMMO in *E. coli* has not been achieved to date as the pMMO genes are reportedly toxic to *E. coli* (Murrell *et al.*, 2000), and aspects such as the expression in ICMs and the presence of multiple copper sites could result in toxicity for microorganisms (Lawton and Rosenzweig, 2016). However, no known studies have attempted to express pMMO in a eukaryotic cell. The multiple compartments which compose eukaryotic cells and the presence of different membrane surfaces could be favourable for expressing the membrane complex and to reduce toxicity. In particular, the endoplasmic reticulum of eukaryotes comprises the largest membranous surface area of the cell, and plays a central role in the synthesis and targeting of proteins destined to the secretory pathway and of integral membrane proteins, housing the synthesis of approximately a third of the proteome (Kriechbaumer and Brandizzi, 2020). Therefore in this work, expression of the pMMO subunits in the ER was evaluated.

The heterologous expression of the pmoC and pmoA subunits from *Methylosinus trichosporium* OB3b fused to fluorescent protein tags was achieved by transient transformation of tobacco (Figure 3.6 and 3.7). The pmoC and pmoA subunits localized to the ER without the addition of any additional targeting signals. It is therefore likely that the proteins are being targeted co-translationally to the ER via the SRP pathway, in an analogous way to the SRP-

mediated translocation mechanism in the plasma membrane in bacteria. In this targeting pathway, the emerging signal peptide or first transmembrane domain at the N-terminus of the protein is recognized by the SRP, and transported to the translocon (Sec61 or SecY in eukaryotes and prokaryotes, respectively) (Rapoport, 2007). In the case of integral membrane proteins, the first transmembrane domain translocates through the channel and inserts into the bilayer by a lateral gate (Rapoport, 2007). For multi-pass membrane protein the process of translocation and insertion continues until the protein achieves its full topology. Interestingly, fusing a fluorescent tag on the N-terminus of pmoC or pmoA did not impair ER targeting. This could be explained as in co-translational targeting the hydrophobic targeting segment must not necessarily reside at the very N-terminus, but must reside at least 65 amino acids from the protein's C-terminus (Hegde and Keenan, 2021). The reason being that the translating polypeptide must be ribosome bound for SRP recognition to occur (Hegde and Keenan, 2021).

Unlike pmoC and pmoA, the pmoB subunit did not localize to the ER in spite of containing an N-terminal signal peptide which directs the protein to the inner membrane in bacteria. GFP-pmoB localized to the cytosol, whereas pmoB-GFP localized to mitochondria with some cytosolic background. Evidently, pmoB's bacterial signal peptide is not recognized by the plant's SRP. This could potentially be due to atypically long N and C regions in the signal peptide, as suggested by bioinformatic studies (Figure 3.31). Cloning signal peptide variants with the shortened N and C regions upstream of Btrunc and a GFP reporter would be needed to formally test this hypothesis but was not attempted here.

One study reported the recombinant expression of the pmoB subunit from *Methylococcus capsulatus* Bath in *E. coli* using a copper tolerant strain (Lu *et al.*, 2019). The subunit fused N-terminally to maltose binding protein was found to localize to the plasma membrane in recombinant *E. coli*. Interestingly, truncated versions of pmoB without its signal peptide were also found to localize to the plasma membrane (Lu *et al.*, 2019). In this study, replacing the bacterial signal peptide by a plant signal peptide, yielding spBtrunc-GFP, allowed to target the subunit to the ER (Figure 3.16). However, the expression levels of this protein were extremely

low. It was thought that this low expression could be due to the protein being targeted downstream from the ER to the apoplast or vacuole. Indeed, spBtrunc-RFP revealed localization in both the ER and vacuole at 2 days after infiltration, and fully vacuolar at 3 days after infiltration. Adding the HDEL C-terminal retention signal to spBtrunc-RFP resulted in the protein having a fully ER localization at both 2 days and 3 days after infiltration, and high expression levels. Surprisingly, adding the HDEL peptide to spBtrunc-GFP only resulted in a marginal increase in expression levels, suggesting that alternative mechanisms must be inducing its downregulation or degradation. The expression of recombinant membrane proteins often yields unpredictable results in terms of expression levels, this could be due to differences in the insertion process in the membrane with the original host, or stress responses arising in the recombinant host, or for unknown reasons (Grisshammer, 2006). Despite the result with spBtrunc-GFP_{hdel}, the spBtrunc-RFP_{hdel} construct resulted in a successful way to produce the pmoB subunit in plant cells.

7.2 Assembly of the pMMO complex in the plant ER

There is a biotechnological interest in MMOs due to their ability to convert methane to methanol at room temperature and pressure. MMOs are complex multi-component enzymes, thus some groups have investigated the creation of functional simpler versions of the enzymes (Lawton and Rosenzweig, 2016). However, there is no current knowledge about a simpler pMMO system that can yield activity, and the assembly of the three subunits as they are found in the methanotroph host would appear to be the strategy with the highest probability of success in a recombinant system. One study examined the functionality of a truncated pmoB subunit without its transmembrane domains, termed soluble pmoB (spmoB), and reported methane oxidation activity of the construct (Balasubramanian *et al.*, 2010). This provided further evidence towards what was thought at the time which was the active site of pMMO being located in the copper site of pmoB. However, this work was later revisited, and it was found

that the methanol production of *spmoB* observed in assays was not actually derived from enzymatic activity (Ross *et al.*, 2019). Moreover, experiments expressing the *pmoB* subunit in *E. coli* did not find enzymatic activity of the subunit (Lu *et al.*, 2019). Most of the recent literature now supports that the active site of the enzyme is the Cu_C site, located in *pmoC*. However, expression of solely *pmoC* is unlikely to yield a functional enzyme, as this copper site, even if coordinated by two histidine residues in *pmoC*, is still located at close proximity with the *pmoA* interface (Figure 7.1 A and B). In addition, the *pmoC* subunits on their own do not interact with each other in the crystal structure (Figure 7.1 C). It is likely that all three subunits are needed to interact and form a specific quaternary structure which has grooves for the access of CH_4 and O_2 and to interact with the electron donor. Further knowledge on the reaction mechanism and electron donor will facilitate in the future guided bioengineering strategies to simplify the structure of the enzyme.

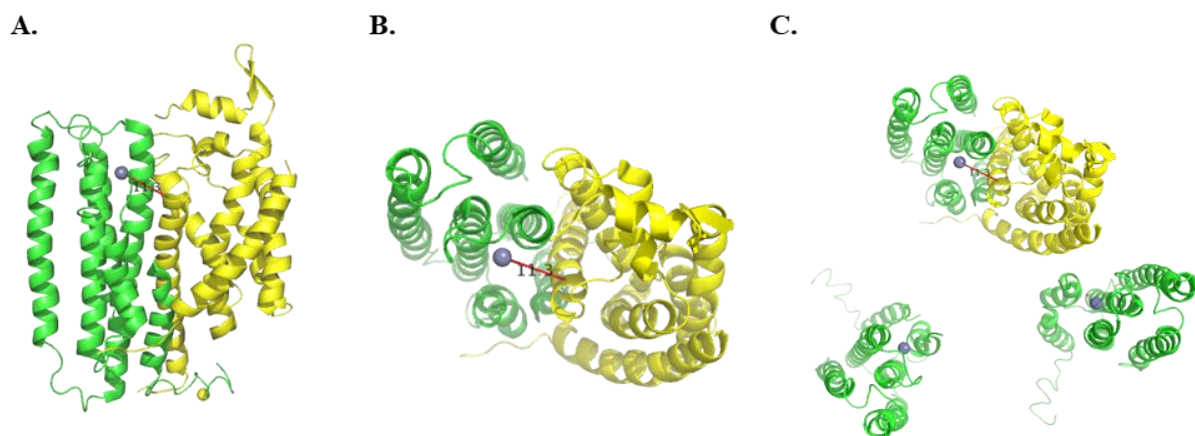


Figure 7.1. Location of the Cu_C site and of the *pmoC* subunits in the nonamer.

The *pmoC* subunit (green) is seen housing a copper site (purple sphere), which is located at a short distance of 11.3 Å from the *pmoA* subunit (yellow), as seen by a side view (A) and a top view (B). The *pmoC* subunits do not interact with each other in the nonamer (C). PDB 3RGB.

In this study pairwise protein-protein interactions were tested via FRET-FLIM as a first line of evidence towards investigating if the pMMO subunits self-assemble to form the nonamer complex on the ER membrane. FRET-FLIM showed that the *pmoC* and *pmoA* subunits

interacted with each other by measuring a significant reduction in the donor lifetime. No interaction was found between spBtrunc and the other subunits. However, this is most likely due to the predicted topology of the proteins, where spBtrunc has both termini in the lumen and pmoC both termini in the cytosol. The occurrence of FRET is highly distance dependent, and no energy transfer resulting in a positive FRET-FLIM signal is expected for proteins located on opposite sides of a membrane (Ratajczak *et al.*, 2014; Gaber *et al.*, 2018). Being guided by the topology in the crystal structure of *Methylosinus trichosporium* OB3b (3CHX), the C-terminus of pmoA was expected to reside in the ER lumen and therefore interact with spBtrunc. However, this was not the case and instead the C-terminus of pmoA interacted with both termini of pmoC, suggesting a cytosolic localization instead. In support of this view, redox-sensitive topology analysis showed both termini of pmoA likely residing in the cytosol. Upon inspection of other crystal structures available in PDB, with significantly higher resolution, the C-terminus of pmoA does appear to face towards the cytosol. Thus, the data in this work supported by that of other crystal structures suggests a topology of both pmoC and pmoA with both termini facing the cytosol and spBtrunc with both termini facing the lumen, and that an interaction occurs between pmoC and pmoA in the plant ER (Figure 3.26).

As an alternative approach to test interactions of spBtrunc, the spBtrunc-RFP construct which has a vacuolar localization at 3 days after infiltration when expressed alone, was co-expressed with clover-pmoA. The co-expression of spBtrunc-RFP with clover-pmoA showed that approximately 50% of cells were partially or fully retained in the ER, instead of being fully vacuolar in 100% of cells at three days after infiltration. This suggested that pmoA and spBtrunc also interact. To confirm an interaction between spBtrunc and the other subunits using FRET-FLIM, a new fusion protein could be constructed placing a fluorophore in the region between the two transmembrane domains of pmoB, which is a small loop facing the cytosol. In addition, native gels could be used to test the presence of the nonamer. However, the evidence presented here provides some indication towards the subunits acquiring the correct topology and being able to self-assemble on the ER membrane.

7.3 Self-cleaving peptides for gene stacking

Early biotechnology projects in plants involved transformation with one or two genes, mainly conferring herbicide or insect resistance to commercial crops (Halpin, 2005). With the accelerated development of molecular biology and with increasing amounts of sequencing data, bioengineering projects now look to insert multiple genes, in order to manipulate whole metabolic pathways, incorporate protein complexes, and study complex traits, amongst others (Naqvi *et al.*, 2010; Ceccon *et al.*, 2020). However, gene stacking is still challenging with low yields of transgenes often occurring unexpectedly, for example due to gene silencing or degradation.

In this work the pMMO subunits were initially co-transformed in a pairwise manner using 2in1 vectors which harbour two expression cassettes within a same T-DNA (Hecker *et al.*, 2015). However, this strategy did not yield expression of both constructs, with some form of silencing occurring between both cassettes when two pMMO subunits were present. This is in contrast to other works which have used the 2in1 vectors successfully to obtain co-expression. For example, Hecker *et al.* (2015) co-expressed the plasma membrane marker SYP121 with the ER marker VAMP723 in colocalization studies, and the nuclear protein BPC6 for FRET (Hecker *et al.*, 2015). Other studies have used 2in1 vectors predominantly for FRET using nuclear localized proteins (Gutsche *et al.*, 2017; Arongaus *et al.*, 2018; Houbaert *et al.*, 2018). It could be that plant marker proteins are more easily co-expressed than exogenous bacterial membrane proteins, but highlights the importance of finding a suitable gene stacking method for each specific group of genes.

A second gene stacking strategy was used to create a polypeptide with 2A self-cleaving peptides located between the pMMO subunits. Self-cleaving peptides are being used increasingly for gene stacking in different cell types, including plants, as they offer many advantages such as being short, having a relatively high cleavage efficiency and high translation efficiency of downstream proteins (Donnelly *et al.*, 2001; Luke *et al.*, 2015).

Through the construction of double fluorophore fusion proteins (RFP-CAB-GFP and RFP-CA-GFP) it was found that self-cleaving peptides allowed the sequential targeting of all three subunits (pmoC, pmoA and spBtrunc) to the ER. The RFP-polypeptide construct was transcribed fully in tobacco, and the most N-terminal protein (RFP-pmoC) and C-terminal protein (clover) were translated following transient transformation of tobacco, and localized correctly to the ER and the cytosol, respectively. It is therefore expected that the pmoA and spBtrunc subunits were also being translated within RFP-polypeptide, as all subunits are translated from a same mRNA. However, some processes may occur such as specific proteolytic degradation or an incomplete cleavage efficiency. Therefore, Western blots could be carried out to confirm translation of all subunits and the cleavage efficiency of each self-cleaving peptide. The fact that the terminal clover localizes entirely to the cytosol, demonstrates that E2A has most likely 100% cleavage efficiency, as its upstream protein (spBtrunc) localizes to the ER. Therefore, even if a fusion protein existed with uncleaved spBtrunc-clover, the chimera would most likely also localize to the ER. The cleavage efficiencies of T2A and P2A could not be established using confocal microscopy as the remaining proteins all localized to the ER. This work, and others in the literature (Luke *et al.*, 2015), suggest that 2A self-cleaving peptides are a promising tool for plant biotechnology, allowing the coordinated co-expression of multiple proteins and targeting to different subcellular compartments.

Even though co-expression and localization of the subunits within polypeptides could be confirmed using microscopy, it was also apparent that the expression levels of polypeptides containing the three subunits (RFP-polypeptide, RFP-CAB and RFP-CAB-GFP) were very low. Frequently, the silencing suppressor P19 is co-expressed transiently to increase transgene yields (Voinnet *et al.*, 2003; Lombardi *et al.*, 2009; Gao *et al.*, 2013; Peyret and Lomonossoff, 2015). In this work, co-expression of the polypeptides with P19 yielded significantly higher expression levels: the intensity level of RFP-pmoC in RFP-polypeptide, increased over three-fold, and the intensity level of clover increased over two-fold (Figure 4.8). Indicating that PTGS mechanisms are responsible for the low expression levels of the polypeptide in tobacco.

Stable T1 Arabidopsis plants containing RFP-polypeptide showed expression of clover in all cells when imaged at the adult stage, whereas no RFP-pmoC was detected. It is intriguing that the clover part of the polypeptide seems to be downregulated to a lesser extent, since PTGS acts upon the whole mRNA, targeting the mRNA with siRNA and marking it for degradation. Therefore, it would be expected for all proteins to be downregulated equally. The RFP-polypeptide produces a long RNA thus it could be possible that the pMMO part of the polypeptide is being targeted for degradation and that the terminal end of the mRNA, even if unstable, could still be translated for some periods of time. Another possible scenario is that additional mechanisms to PTGS are occurring, such as targeted protein degradation of the pMMO subunits and not of clover, producing this differential expression pattern. 5-day old T2 Arabidopsis seedlings did show trace levels of expression of RFP-pmoC alongside expression of clover, indicating that the downregulation begins at a very early growth stage. Whilst these young seedlings could be used in enzymatic assays to test activity, producing stable plants with RFP-polypeptide and a silencing suppressor might be a more viable strategy to yield sufficient levels of product for its detection.

7.4 Impact of the single subunits and the pMMO polypeptide on ER structure and dynamics

The incorporation of transgenes can often result in toxicity to the host cell or in serious modification to the cellular environment. In this case, transient expression of pmoC in tobacco resulted in disruption of the ER network in all cells at 2 days after infiltration, and in visible necrosis of the leaf tissue (both at OD₆₀₀ 0.1 and 0.05) (Figure 3.8). It is unknown what aspect of the protein induces this phenotype. On the other hand, the pmoA and spBtrunc subunits do not cause such necrosis, and a varying number of cells can express pmoA and spBtrunc-RFP at high levels maintaining a contiguous and seemingly normal ER structure by visual inspection. Any co-expression of pmoC with another subunit from individual plasmids, or in

the RFP-CA-GFP or RFP-CA polypeptides, also results in disruption of the ER in all cells, presumably induced by the presence of pmoC. Surprisingly, expression of the RFP-polypeptide construct to high levels with P19, which includes pmoC alongside pmoA and spBtrunc, does not produce disruption of the ER in all cells, and instead produces seemingly normal ER in many cells. This suggests that the pmoC subunit may form quaternary structures with pmoA and spBtrunc which are considerably less detrimental to the cell than when it is expressed alone or with a single other subunit. In the case that the pMMO complex were to be assembling correctly from RFP-polypeptide (which still requires confirming with native gels), this gives a hopeful result that plant cells are able to harbor the pMMO complex without significant damage, as opposed to what has been found so far in the literature regarding attempts to express the complex in *E. coli*.

Besides from overall changes in the levels of ER disruption, other more subtle alterations can occur such as changes in the size or shape of cisternae or tubules, a shift towards a higher proportion of tubules or cisternae, or a change in ER dynamics. The ER morphology is thought to be connected to its secretory capacity, and is seen to vary according to the plant's developmental stage, in response to biotic stresses such as viral infection (Griffing *et al.*, 2016), and in response to drug treatments and protein over-expression (Pain *et al.*, 2019). The ER plays a central role in many cellular functions such as protein and lipid biosynthesis, protein folding and secretion, calcium homeostasis, amongst others. Moreover, it presents contact sites with multiple organelles (Griffing *et al.*, 2016). Therefore, maintaining ER homeostasis in the recombinant system is important to avoid perturbing other cellular processes. Some alterations in the ER are not easily identified by visual inspection due to the natural cell to cell variability that already exists in ER morphology and movement within a same tissue (Pain *et al.*, 2019). In this work the AnalyzER software (Pain *et al.*, 2019) was used to quantify and statistically compare several parameters of ER structure and movement. The RFP-tagged subunits and RFP-polypeptide were co-expressed with the luminal marker GFP-HDEL and compared with GFP-HDEL alone (considered “wild type” ER). The luminal channel was analyzed in all cases

for consistency in the analysis. Upon co-expression of pmoA with GFP-HDEL, it became notable that the lumen was highly modified. A phenotype was observed similar to that of reticulon proteins which cause constriction of the ER, consisting of a marked thinning of tubules and the appearance of discrete bulges in the luminal content (Sparkes *et al.*, 2010; Breeze *et al.*, 2016). A consistent skeleton of the ER network using the luminal channel of pmoA could not be obtained, for this reason pmoA expressed alone was excluded from the subsequent software analysis.

Boxplots did not show large differences in the morphology of cisternae or polygonal regions between RFP-polypeptide and the GFP-HDEL control, or between spBtrunc-RFPhdel and GFP-HDEL. Amongst the cisternal texture metrics, it was observed that RFP-polypeptide and spBtrunc-RFPhdel had a lower correlation than the control, reflecting less homogeneous cisternae. RFP-polypeptide also showed bulges in the luminal content along some tubules. Trace analysis was carried out to detect the presence of a “reticulon like” constriction phenotype, indicated by a negative correlation between the peaks and troughs of both channels. However, a clear constriction phenotype such as that seen using trace analysis in RTN1 (Pain *et al.*, 2019) was not observed. Some peaks in the green channel did correspond with a trough in the red channel, which could be induced by constriction, however, other peaks in the green channel coincided with peaks in the red channel, or with a constant intensity in the red channel, rather than presenting a constant overall negative correlation predicted from the reticulon studies (Breeze *et al.*, 2016; Pain *et al.*, 2019).

AnalyzER was also used to study ER dynamics using an optical flow method based on the Farnebäck algorithm (Farnebäck, 2003), and to study changes in static (or persistent) elements. Boxplots did not show large differences in the average speed or maximum speed of tubules, and showed a fractionally lower average speed and maximum speed of cisternae in spBtrunc-RFPhdel. Interestingly, both the tubule and cisternal speeds of the GFP-HDEL control in this work were lower than those determined previously for untreated cells (Pain *et al.*, 2019). This is probably due to the plants being grown in a different year and season, which highlights the

importance of taking the full dataset for analysis in the same time frame. Similarly to the trends found with the speed, the mean persistency of tubules did not show great differences between the groups, and spBtrunc-RFPhdel showed a slightly higher mean cisternal persistency.

A MANOVA analysis was carried out using 19 parameters of morphology and movement (Pain *et al.*, 2019) to study the impact of the polypeptide or spBtrunc upon the ER as a whole. MANOVA showed that there are significant differences between the groups. In the accompanying clustering analysis based on the multivariate means the RFP-polypeptide appears to be more similar to GFP-HDEL than spBtrunc-RFPhdel. Post-hoc comparisons showed very few metrics presenting significant differences. The only metric which differed between GFP-HDEL and RFP-polypeptide was the cisternal correlation, and three metrics differed between GFP-HDEL and spBtrunc-RFPhdel (cisternal correlation, cisternal speed and persistency). The differences were markedly less than those found between the GFP-HDEL control and drug treatments, heat shock, or over-expression of LNP2, where most of the variables included in the analysis differed significantly from the control (Pain *et al.*, 2019). Overall, this analysis suggests that the RFP-polypeptide does not induce considerable changes to ER structure and dynamics.

7.5 Towards creating pMMO rice

7.5.1 Physiological aspects: reducing equivalents, copper requirement and methanol production

Experiments indicated that tobacco and Arabidopsis plants are able to produce the pMMO polypeptide, without major changes to ER structure and dynamics, however its expression is highly downregulated. It was not determined in this work whether the polypeptide is functional due to restrictions deriving from the COVID-19 pandemic, however, the description of enzyme assays is presented (Chapter 6). The experimental design includes propylene epoxidation assays to be carried out with and without the addition of duroquinol to act as a reducing agent.

Importantly, if pMMO were to be active in the plant lysates but only with the addition of duroquinol, this would indicate that no native components in the ER membrane are able to reduce the recombinant enzyme. The native electron donor of pMMO has not been established with certainty, but if the enzyme can only be reduced with quinols a feasible strategy could be to introduce the pMMO polypeptide in thylakoid membranes or in the mitochondrial inner membrane where plastoquinone and ubiquinone pools are present, respectively (Liu and Lu, 2016). Alternatively, further bioengineering strategies could be devised to incorporate a suitable reductase component in the ER. This would seem a challenging task since reductases present in the ER membrane are specific to the protein they reduce (such as cytochrome P450 reductase and cytochrome b5 reductase). Perhaps further knowledge on the pMMO reaction mechanism and comparison with other similar enzymes could help to identify a reduction method which is independent of quinols and could then be engineered in the ER.

The experimental design of enzyme assays also includes a test of whether the supplementation of copper is needed for activity. pMMO has at least two monocopper centres per promoter, therefore at least 6 copper ions per nonamer, although some studies have reported higher copper levels. In plants copper is an essential micronutrient, as it is present as a cofactor in enzymes required for photosynthesis, respiration, ethylene sensing, and oxidative stress response, amongst others (Yruela, 2009). Certain concentrations of the micronutrient are needed, however, excess Cu can cause toxicity as it can produce reactive oxygen species (Yruela, 2009). Copper transporters are present in the plasma membrane to uptake the metal, and in other organelles such as the ER. In particular, the ER contains the RAN1 copper transporter, which supplies copper to the membrane-bound ethylene receptor (Yruela, 2009). Thus, it is possible that sufficient copper ions would be available in the ER membrane for pMMO, depending on expression levels. In the case of requiring higher copper levels to support high pMMO expression, the use of copper tolerant rice strains could be considered (Liu *et al.*, 2008).

Another factor to consider in the case of obtaining enzymatic activity is the resultant production of methanol, which could potentially be harmful for the plant. Plants do not have a known

metabolic pathway for the assimilation of methanol due to the absence of enzymes known to catalyze methanol (such as methanol dehydrogenase in bacteria). Methanol is produced naturally in plants, principally due to the degradation of pectin methyl esters in the cell wall (McGiffen and Manthey, 1996; Yurimoto *et al.*, 2021). The volatile compound exits the leaves through the stomata, and methanol consuming methylotrophs are known to inhabit the leaf surfaces consuming part of the alcohol, with the remaining part being released to the atmosphere (McGiffen and Manthey, 1996; Yurimoto *et al.*, 2021). Some early studies have reported an enhancement in growth yield in several C₃ species with the application of methanol sprays, however, these studies were not reproduced by other groups and the biochemical pathway which could explain this growth enhancement was not found (McGiffen and Manthey, 1996). Even though an enzymatic pathway for the conversion of methanol in plants has not been uncovered, studies using labelled methanol have found that the alcohol is oxidized to CO₂ in several plant species (Cossins, 1964) and assimilated into amino acids and other organic molecules (Cossins, 1964; Gout *et al.*, 2000). It would also appear that plants are resilient to certain levels of methanol. For example, 5mM methanol did not affect growth parameters in sycamore cells (Gout *et al.*, 2000), and applying up to 10% v/v of methanol solution as a foliar spray in Arabidopsis, tobacco and tomato did not lead to phytotoxic damage (Ramírez *et al.*, 2006). However, the application of methanol solution to the roots of tobacco and Arabidopsis did induce phytotoxic damage (Ramírez *et al.*, 2006).

It is challenging to predict how much methanol would be produced in rice containing pMMO, as the values reported in the literature for methane emissions per plant are highly variable and they are also dependent on the growth stage and geographic conditions. However, if the plants were to consume most of the methane being transported to the atmosphere, a considerable amount of methanol would be produced. Experiments could be designed to test at which point methanol production becomes toxic. For example, once pMMO is designed to work at the maximum specific activity possible in the plant, mutants could be created with different expression levels of pMMO, i.e. under promoters with different strengths, or by expressing

solely in the root or in the aboveground biomass, as methanol production in the stems and leaves could be less toxic than in roots. Also further bioengineering which promotes the growth of methanol consuming methylotrophs on the leaves of pMMO containing rice could be explored.

7.5.2 Obstacles and successes in the development of GM rice

Whilst optimising enzymatic activity of pMMO is easier in the model plants tobacco and *Arabidopsis*, the strategy can then be transferred to expressing functional pMMO in rice plants. Previous biotechnology initiatives in rice have created transgenic varieties for diverse purposes, for example to produce plants that are more resistant to drought and salinity, to confer resistance to biotic stresses, or for biofortification, amongst others (Biswal *et al.*, 2017). A remarkable case of rice bioengineering was the creation of Golden Rice, which produces a grain with high beta-carotene content aimed at combating vitamin A deficiency in developing countries (Ye *et al.*, 2000). The creation of Golden Rice involved co-transformation with two recombinant genes: the gene encoding phytoene synthase (*psy*) and the gene encoding carotene desaturase (*crtI*) (Ye *et al.*, 2000; Paine *et al.*, 2005), incorporated by co-transformation within a same T-DNA region. Another notable example of transformation with multiple genes was the creation of biofortified “Purple Endosperm Rice”, containing high contents of anthocyanin and antioxidant activity (Zhu *et al.*, 2017). Purple Endosperm Rice was achieved by the transformation with eight transgenes (two regulatory genes and six structural genes), and the activation of 13 endogenous genes (Zhu *et al.*, 2017). Currently a major bioengineering initiative taking place is the multi-institutional C₄ rice project, which aims to transfer C₄ photosynthesis to the crop to improve its photosynthetic efficiency and to increase yields (www.c4rice.com).

New GM crops will most likely play a major role in the future of agriculture to help provide the required increase in productivity to produce sufficient food to our growing population, especially in a changing environment, with higher CO₂ levels and higher temperatures

producing drought, higher salinity, floods, amongst others. Once the research and field trial phases are completed, approving a GM crop for commercialization is still a very difficult task, and opposition from the general public in many countries is still strong. For example Golden Rice 2, which was published in 2005 (Paine *et al.*, 2005) has only been approved for commercialization in 2017/2018 in a few countries (Australia, New Zealand, Canada and the US) (Ceccon *et al.*, 2020). Besides from the lengthy approval process, incidents occurred during field trials, such as the vandalization of a Golden Rice field in the Philippines in 2013 by protesters. However, the fact that the modified crop reached approval in some countries offers hopeful outcomes to further biotechnology initiatives. The creation of biofortified rice requires expression of the transgenes in the rice grains. “pMMO rice” could be produced using a promoter specific to roots and stems which does not produce expression in the rice grain, due to methane being transported in the plant almost entirely through the roots and stems. Therefore, even though the genes will be present throughout the plant, the fact that no transgene expression occurs in the edible part of the plant could somehow facilitate its application.

8 Future work

8.1 Testing enzymatic activity of RFP-poly and RFP-polyM *in vitro*

The most immediate future work will need to test enzymatic activity of the plant lysates obtained after transient expression of RFP-poly and RFP-polyM with p19. This would be carried out as described in Chapter 6, via a propylene epoxidation assay to detect propylene oxide with gas chromatography (Kitmitto *et al.*, 2005). The assays would be carried out with and without supplemental copper added to the lysis buffer, and with and without the addition of duroquinol to the assay medium. Besides from duroquinol, NADH and NADPH could also be supplemented to the assay medium as plant reductases in the ER membrane use these molecules as reductants (such as cytochrome P450 reductase and cytochrome b5 reductase which use NADPH and NADH, respectively). The subsequent steps would be dependent on the results obtained by assays.

In the case of obtaining pMMO activity in any samples containing NADPH and NADH, this would indicate that the enzyme is active on the ER membrane, therefore, further assays could be carried out *in vivo* to confirm this result in a more physiological condition, and methane oxidation could also be tested *in vivo* (see section 8.2). In the case of obtaining pMMO activity solely with duroquinol, further bioengineering strategies should be investigated to add a reductase component for pMMO on the ER membrane. If the investigation does not yield a plausible strategy to add a reductase, expression of the pMMO polypeptide in the thylakoid or inner mitochondrial membrane, which contain quinones, could be attempted (see section 8.3).

Finally, if no activity was found *in vitro*, the reason for this outcome could be investigated. Firstly, it should be confirmed that the recombinant proteins are present in the plant lysates using Western blots. Even though protein expression is confirmed in the leaf pieces by confocal microscopy, it is possible that proteins are lost during the extraction, or that the microsomal

fractions containing pMMO do not partition to the supernatant properly during the preparation of the lysates. Western blots could be carried out with antibodies against the 2A self-cleaving peptides, as antibodies exist which target the conserved C-terminal region of different 2A self-cleaving peptides which remains attached to the upstream protein. All three subunits could be distinguished by size with this method. Further to establishing the presence of the proteins, it would be favourable to carry out native gels to determine if pMMO is assembling correctly on the ER membrane as a nonamer and if can be detected as a nonamer in the plant lysates. If the quaternary structure disassembles during the preparation of lysates this could be a reason for loss of activity *in vitro* and further optimization of the extraction conditions would be needed. In parallel, propylene epoxidation assays could be carried out *in vivo* to confirm that the absence of a signal is not due to the extraction procedure.

8.2 Performing enzyme assays *in vivo*

The advantage of *in vitro* assays is that they are potentially more sensitive and could facilitate product formation due to the enzyme being more directly in contact with the substrates, and being able to supplement the assay with other components. However, *in vivo* assays pose the advantage of not requiring an extraction step, which could lead to inactivation of the enzyme, and being able to test the system in physiological conditions. In this work stable Arabidopsis plants containing RFP-poly were created, however, they had trace expression of RFP-pmoC and only at a very young age. This agreed with results from transient transformation of tobacco which required the co-expression with p19 to enhance expression levels, suggesting a downregulation via PTGS. Therefore, the creation of stable plants which include a silencing suppressor would be highly recommended. Previous studies have created Arabidopsis PTGS mutants (with mutated *sgs2* and *sgs3*) (Mourrain *et al.*, 2000), and these have been shown to greatly enhance transgene expression, as well as showing no phenotypic differences with wild type plants (Butaye *et al.*, 2004). It would therefore be a suitable strategy so transform RFP-

poly and RFP-polyM into the Arabidopsis PTGS mutant background in order to perform *in vivo* assays on seedlings which express pMMO at high levels, to test whether pMMO can detoxify methane in plants. The PTGS mutants showed similar phenotypes to the wild type in contained growth conditions (Butaye *et al.*, 2004), it is likely that in field conditions they would be more susceptible to viral and bacterial infections and therefore not pose a viable strategy for the final application. In this case, other strategies could be explored that would reduce the PTGS specifically for this transgene and not systemically in the plant.

8.3 Expression of pMMO in the chloroplast

8.3.1 Advantages of chloroplast transformation

As an alternative approach to expression in the ER, the expression of pMMO in the chloroplast could be investigated. In particular, the thylakoid membranes have a plastoquinol pool which could act as reducing equivalents for the recombinant enzyme. Moreover, the chloroplast also contains copper transporters, and the abundant copper protein plastocyanin, located in the thylakoid lumen (Yruela, 2009), therefore copper levels should be sufficient for pMMO. However, targeting RFP-poly to the chloroplast could potentially be challenging. Most chloroplast proteins are encoded in the nuclear genome, and are imported to the organelle through the TOC and TIC complexes (in the outer and inner chloroplast envelopes, respectively) (Jarvis, 2008). Proteins are predominantly targeted to the chloroplast post-translationally, by means of an N-terminal transit peptide, and a second targeting signal directs proteins to the thylakoid membrane and lumen (Jarvis, 2008). In RFP-poly, if a transit peptide were to be added to the N-terminal protein, this should be targeted correctly to the chloroplast, but the downstream proteins would not, due to targeting being post-translational. Thus, each of the three subunits would need to be cloned individually with a transit peptide and correctly targeted to the chloroplast individually.

Alternatively, direct chloroplast transformation can be attempted for the transgenes to integrate in the chloroplast genome. A notable benefit of chloroplast transformation is the possibility of transforming with full operons, due to the prokaryotic nature of the organelle (Verma and Daniell, 2007). Chloroplast transformation has been exploited in biotechnology due to its many advantages: the ability to achieve gene stacking through operons in a single transformation event, obtaining high protein levels due to the large number of chloroplasts present in the cell, allowing higher transgene containment due to maternal inheritance, and enabling site-specific recombination through the use of flanking homologous regions (Verma and Daniell, 2007). Several cases exist of successful multi-gene chloroplast transformation, such as the introduction of the bacterial polyhydroxybutyrate (PHB) operon from bacteria in tobacco, containing three genes which encode the synthesis of the bioplastic. Although initial attempts yielded plants with male sterility and growth deficiency (Lössl *et al.*, 2003), further optimization overcame these issues and yielded healthy plants with a high content of PHB (17.3% dry weight PHB in leaves) (Bohmert-Tatarev *et al.*, 2011). Other examples include the expression of the *cry2Aa2* operon in tobacco to produce insecticidal compounds, accumulating 45.3% of total soluble protein in leaves (De Cosa *et al.*, 2001), and complex endeavors such as the incorporation of the artemisinic acid biosynthesis pathway, achieved via transformation with 12 genes from different organisms (Saxena *et al.*, 2014).

8.3.2 Localizing transgenes to the thylakoid membrane

The chloroplast is a complex organelle with several internal compartments such as the outer and inner envelopes, the stroma, the stacked thylakoid membranes and the thylakoid lumen. Ideally chloroplast-encoded pMMO would be localized to the thylakoid membrane to obtain electrons from the photosynthetic electron transport chain. A previous study has recombinantly expressed membrane proteins in the chloroplast which used photosynthetic reducing power to perform their function. In that study two ER membrane-bound P450 monooxygenases and a cytosolic protein involved in the biosynthesis of dhurrin were expressed in tobacco chloroplasts

using chloroplast transformation (Gnanasekaran *et al.*, 2016). The P450s which are normally reduced by ER-bound P450 reductase, used photosynthetic reducing power in the form of reduced ferredoxin. The authors showed that the P450s were expressed in the chloroplast membrane fraction, stating that this would be primarily the thylakoid membrane, but could not discount that a fraction of the proteins were found in the inner envelope (Gnanasekaran *et al.*, 2016).

Although the targeting of nuclear encoded proteins to chloroplast compartments has been extensively studied (Jarvis, 2008), the targeting of chloroplast encoded proteins has not been characterized in such detail. An SRP pathway or spontaneous insertion are two routes whereby proteins with a transit peptide may then be targeted to the thylakoid membrane once inside the chloroplast (Jarvis, 2008), and could also be pathways through which chloroplast encoded proteins insert into the thylakoid membrane. One study recombinantly expressed a cyanobacterial bicarbonate transporter, containing 12 TM domains, through chloroplast transformation in tobacco, and found that approximately 25% of the transgenic protein localized to the inner envelope, and 75% localized to the thylakoid membranes (Pengelly *et al.*, 2014). It is likely thus, that a significant proportion of the pMMO subunits would be localized to the thylakoid membranes if expressed from the chloroplast genome. This could be tested by fractionation of chloroplast membranes and Western blots.

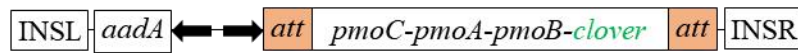
8.3.3 Preliminary work on the chloroplast transformation with pMMO

In this project, preliminary work was carried out towards transforming chloroplasts with pMMO as an alternative strategy for robust expression. The full pMMO operon sequence from *Methylosinus trichosporium* OB3b was cloned from genomic DNA, fused to a clover fluorophore and placed into a chloroplast transformation vector (Gottschamel *et al.*, 2013) by Dr Verena Kriechbaumer (Figure 8.1 A). Transformation was performed via particle bombardment with the help of Dr Frances Tolmie, and selection of leaf pieces following bombardment was carried out in MS media containing 500 mgL⁻¹ spectinomycin and hormones

for the regeneration of shoots. Two shoots were recovered following bombardment yielding plants “A” and “B” (Figure 8.1 D). In the first round of selection plants are heteroplasmic (some chloroplasts are transgenic and some are not). A second round of selection starting from a leaf piece of the heteroplasmic plant is needed to achieve homoplasmy (Verma *et al.*, 2008). In this preliminary work, two plants were regenerated from leaf pieces from each A and B plant yielding “subclones”. The seeds of these subclones were resistant to spectinomycin (Figure 8.2). The subclones of plant A appear to be homoplasmic, as they have a near 100% germination proportion (Figure 8.2 C, D), whereas the subclones of plant B appear to still be heteroplasmic as they have low germination proportions (Figure 8.2 E, F).

Future work involves genotyping these plants to confirm the presence of the transgenes, as well as confirming homoplasmy and chloroplast integration with PCR. Subsequently, RT-PCR could be used to test transcription, and finally Western blots and enzyme assays to test protein production and activity. The disadvantage of this method is that the transformation and regeneration procedure is very lengthy (it takes more than 6 months to obtain homoplasmic plants), and expertise in routine chloroplast transformation would be favourable. For example, in this study, several additional rounds of bombardment were performed with the pMMO sequence without clover, and with the original chloroplast transformation vector containing GFP within the Gateway recombination sites as a control. However, no viable plants were regenerated. Overall, the frequency of a successful regeneration event following bombardment was very low compared to that reported by other groups.

A. pEXPN-pMMO-clover



B.

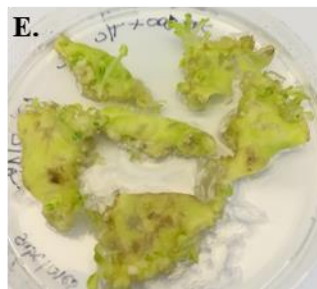
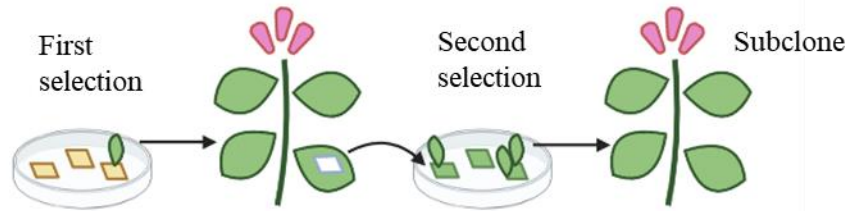


Figure 8.1. Chloroplast transformation of the pMMO operon.

(A) Simplified diagram of the chloroplast transformation vector (Gottschamel *et al.*, 2013) containing the pMMO operon with a clover fluorophore, showing the left and right homologous recombination regions (INSL and INSR), the spectinomycin resistance gene (*aadA*), and Gateway cloning sites (*att*). Black arrows indicate promoter sequences. (B) Diagram of the regeneration procedure following particle bombardment. During the first round of selection, leaf pieces were placed in regeneration medium containing antibiotics and turned a pale colour if not transformed, whereas a shoot developed from a cell which was putatively transformed. The shoots were grown to maturity and a leaf section from the grown plant was placed once more in selection media to achieve homoplasmy. The plants grown from the shoots during the second round of selection were termed subclones. (C) Photo of a shoot obtained during the first round of selection. (D) Two mature plants (plants “A” and “B”) originated from two shooting events during the first round of selection. (E) Second round of selection.

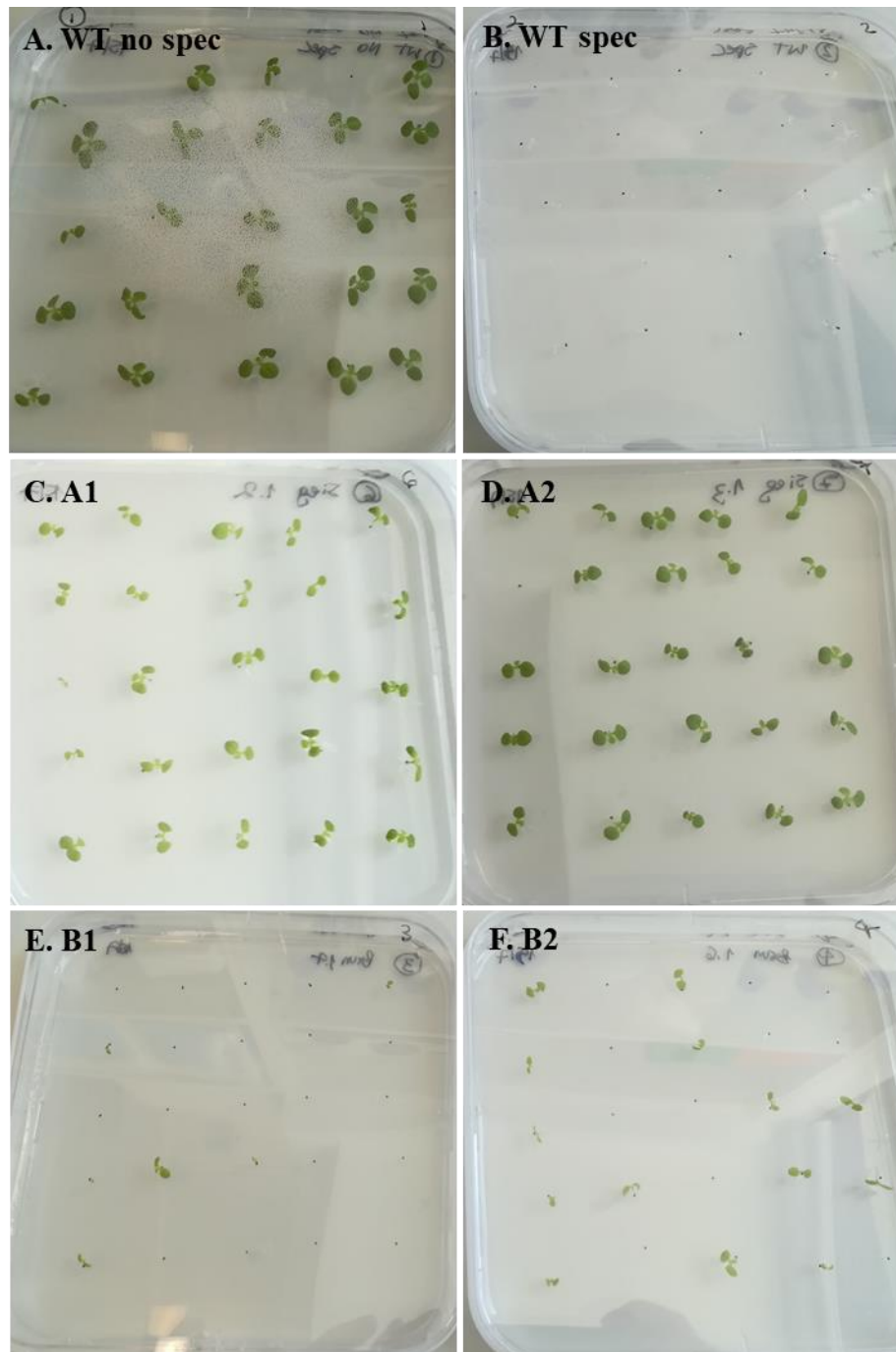


Figure 8.2. Selection on plates following chloroplast transformation.

2-week old tobacco seedlings on MS media without antibiotics (A), and on MS media with 500 mgL⁻¹ spectinomycin (B-F). (A and B) WT seeds, (C and D) subclones from plant A, (E and F) subclones from plant B.

9 Supplementary data

9.1 Chapter 2 – Materials and methods

Appendix 1: Codon optimized polypeptide sequence

Pmoc-**T2A**-pmoA-**P2A**-spBtrunc-**E2A**-clover

>polypeptide nucleotide sequence

ATGTCTGTCACAACTGAAACAACCTGCTGGAGCGGCAGCAGGATCAGACGCGATTGTTGA
CCTTCGGGGCATGTGGGTTGGTGTTCGGGATTGAACATTTTTATCTGATCGTTTCGCATA
TATGAGCAGATATACGGGTGGAGGGCCGGTTTGGATTTCGTTCCGCCCTGAGTTCCAGACA
TACTGGCTCAGTATACTCTGGACCGAGATTCCGTTGGAGTTGGTTTCAGGCTTGGCGCTG
GCCGTTGGTTGTGGAACACCAGAGACAGGAACGTTGACGCCGTGGCACCTAGGGAAGA
GCTCAGGAGACATGTGGTGCTTGTGGAGTGGCTCGTCGTTTACGCCGTGGCTATTTACTG
GGGCGCCTCCTTCTTCACTGAACAAGATGGGACCTGGCACATGACCGTGATACGCGACA
CAGATTTTACCCCGAGTCATATTATCGAGTTTACATGTCATACCCCATCTATAGTATCAT
GGCGGTGGGAGCGTTCTTCTACGCAAAGACAAGGATCCCGTATTTCCGCGCACGGCTTCTC
CTTGGCCTTCCTGATAGTAGCGATCGGCCCTTCATGATCATCCCAAATGTCGGTCTCAAT
GAATGGGGCCATACGTTTTGGTTTATGGAGGAGCTGTTTCGTGGCTCCACTCCATTGGGGG
TTCGTCTTCTTCGGCTGGATGGCCCTGGGAGTATTCGGCGTCGTCCTGCAGATCCTCATGG
GTGTCAAGCGTTTGATTGGTAAGGATTGCGTTGCCGCGCTTGTGGGGGAAGTGGCGAAG
GACGCGGTTCTTCTGACGTGTGGTGATGTGGAGGAGAATCCCGGTCCGATGTTTCACAA
GCAAGAGTGGTGGGGCTATCGGTCCATTTCACTCCGTCGCTGAGGCTGCTGGCTGTGTGA
AGACTACGGACTGGATGTTCTTGACACTCCTCTTCTTGGCGGTATTGGGGGGGTACCATA
TTCATTCATGTTGACGGCCGGAGACTGGGACTTTTGGGTGGATTGGAAGGATAGGCGTA
TGTGGCCTACAGTGGTCCCAATCCTGGGTGTGACTTTTGCAGCCGCGGCGCAAGCCTTTT
TCTGGGAGAATTTCAAACCTCCCTTTTCGGAGCAACCTTCGCGGTTTCGGGCCTGCTTATCG
GCGAGTGGATCAACCGCTACTGCAACTTTTGGGGATGGACCTACTTCCCAATTTCACTGG
TGTTCCCGAGCGCTTTGGTGGTTCCAGCACTGTGGCTCGACATAATCATGCTTCTCTCAGG
ATCTTACGTAATCACTGCAGTGGTCCGGTCCCTTGGATGGGGTCTTCTCTTACCCTAAT
AACTGGCCAGCTATTGCCGCACTGCACCAAGCGACCGAACAACACGGACAACCTTATGAG
CCTCGCAGACCTCGTGGGGTTTCACTTCGTCCGGACCTCTATGCCGGAATATATCCGCAT
GGTTGAGCGCGGAACATTGCGGACTTTCGGTAAGGAGGTGGTCCCGGTTGCAGCCTTCTT
TTCTGGCTTTGTGTCGATGATGGTATACTTCTCTGCTGGTTCGTGGGGAAATGGTATAGC
ACCACCAAAGTGATACAAAAAATAGGTTTCGGGGGCTACGAACCTTCAGTCTGCTGAAACA
GGCGGGAGACGTTGAGGAGAATCCTGGTCCAATGGCTCGCCTTACCTCGATTATCGCCCT
TTTCGCCGTGCGACTGCTGGTTGCAGATGCGTATGCCTACAGAACGCACGGCGAGAAATC
GCAGCAGGCGTTTCTCCGCATGCGCACTCTCAACTGGTACGATGTAAATGGTCTAAGAC
ATCCCTTAATGTAAATGAATCGATGGTCCTTCTGGCAAGGTACACGTCTTTTCGGCGTGG
CCCCAGGCAGTTGCGAATCCTAAAAGCTCCTTCTCAACGCCGCGGAACCTGGCCCTGTT
CTTGTTTCGTACAGCTCAGTTCATCGGGGAGCAGTTTGGCCCTAGGTCAGTGTCAATTGGAG
GTGGGAAAAGATTACGCGTTCTCGATAGATCTTAAGGCGCGTAGGGCCGGTAGGTGGCA
TGTTACGCTCAGATTAACGTAGAGGGCGGAGGTCCTATCATTGGCCCTGGACAGTGGAT

TGAAATCAAGGGAGACATGGCAGACTTCAAAGATCCAGTCACATTGCTGGACGGGACAA
 CCGTCGATCTCGAGACTTACGGAATCGATCGGATCTACGCCTGGCATTTCATGGATGA
 TAGCTGCTGCCGCGTGGATTCTCTATTGGTTCTTTAAAAAGGGGATCATTGCCTCGTACCT
 GCGCATATCGGAAGGGAAGGATGAAGAGCAGATTGGTGACGACGACCGGAGGGTCGGT
 GCAATTGTCCTTGCTGTGACAATTTTGGCGACTATCATAGGATACGCTGTGACAAATTCA
 ACCTTCCCAAGGACAATTCCTCTGCAGGCTGGACTCCAGAAGCCTTTGACGCCGATTATA
 GAGGAAGGGACGGCGGGAGTTGGTCCGCATGTGGTTACGGCAGAGCTCAAAGGGGGTGT
 ATACAAAGTCCCGGGCAGAGAGCTGACGATACAAGTTAAGGTTACCAATAAGACTGATG
 AGCCCCTCAAGCTCGGAGAGTATACAGCCGCGGGATTGCGGTTTCTCAATCCTGATGTAT
 TCACGACGAAGCCCGAGTTTCCAGACTACCTCTTGCGGACAGAGGACTTTCTACGGATC
 CTACCCCACTGGCTCCAGGCGAGACGAAAACGATTGAGGTAAAAGTTCAAGACGCTCGC
 TGGGATATTGAGCGTCTGTCTGGATCTGGCTTACGACACAGACAGCCAAATAGGCGGCCT
 CCTGATGTTTTTCTCCCCCTCCGGAAGCGGTACGCCACCGAAATTGGCGGCCCGGTAAT
 TCCGAAATTCGTGGCTGGAGACATGCCTGGGTCAGGGCAGTGCACGAACTATGCACTCTT
 GAAACTGGCTGGAGATGTTGAGTCTAATCCAGGCCCGATGAGCAAGGGCGAGGAGCTTT
 TTACCGGGGTGGTGCCGATTCTCGTGGAAGCTTGATGGGGACGTCAACGGCCATAAGTTCA
 GTGTGTCGGGCGAAGGGGAAGGAGATGCGACCTATGGCAAAGTACCCCTCAAGTTCATA
 TGCACCACGGGCAAAGTCCCGTTCCATGGCCGACTCTCGTCACGACTTTCTCATATGGC
 GTCCAATGCTTCAGTCGGTACCCCGACCACATGAAGAGACACGACTTCTTCAAGTCGGCC
 ATGCCTGAGGGATATGTCCAGGAGCGCACAATTTTTTTTAAAGATGATGGCAACTACAAG
 ACACGCGCCGAAGTTAAGTTCGAGGGCGACACGCTTGTCAATCGTATCGAATTGAAGGG
 CATTGATTTTAAAGGAAGACGGGAATATCTTGGGCCACAAGCTGGAGTATAATTACAATA
 GCCACAACGTGTATATAATGGCGGACAAGCAGAAGAATGGCATCAAAGCCAACTTTAAA
 ACACGCCACAACATCGAGGACGGCGGGGTCCAAGTCTGATCACTACCAGCAGAACAC
 TCCAATTGGGGATGGTCCGGTTCTCCTTCCTGATAACCATTATCTGAGTACACAGTCAGC
 CCTCAGTAAGGACCCGAATGAAAAGCGCGATCATATGGTGTGCTCGAATTCGTTACAGC
 CGCCGGAATAACCCACGGCATGGATGAGCTCTATAAGTAG

>polypeptide protein sequence

MSVTTETTAGAAAGSDAIVDLRGMWVGVAAGLNIFYLVIRIYEQIYGWRAGLDSFAPEFQTY
 WLSILWTEIPELVSGLALAGWLWKTRDRNVDAVAPREELRRHVVLVEWLVVYAVAIYWG
 ASFFTEQDGTWHMTVIRDTDFTPSHIIEFYMSYPIYSIMAVGAFFYAKTRIPYFAHGFSLAFLIV
 AIGPFMIIPNVGLNEWGHTFWFMEELFVAPLHWGFVFFGWMALGVFGVVLQILMGVKRLIG
 KDCVAALVGGSGEGRGSLTTCGDVEENPGPMFTSKSGGAIGPFHVAEAAAGCVKTTDWMFL
 TLLFLAVLGGYHIHFMLTAGDWDWFDWVDRRMWPTVVPILGVTFAAAAQAFFWENFKLP
 FGATFAVSGLLIGEWINRYCNFWGWTYFPISLVFPSALVVPALWLDIIMLLSGSYVITAVVGS
 LGWGLLFYPNNWPAIAALHQATEQHGQLMSLADLVGFHFVRTSMPEYIRMVERGTLRTFGK
 EVVPVAAFFSGFVSMMVYFLWWFVGKWySTTKVIQKIGSGATNFSLLKQAGDVEENPGPM
 ARLTSIIALFAVALLVADAYAYRTHGEKSQQAFLRMRTLNWYDVKWSKTSNLNVNESMVL
 SKVHVFSAWPQAVANPKSSFLNAGEPGPVLVRTAQFIGEQFAPRSVSLEVVKDYAFSIDLKA
 RRAGRWHVHAQINVEGGGPIIGPGQWIEIKGDMADFKDPVTLLDGTTVDLETYIDRIYAWH
 FPWMIAAAAWILYWFFKKGIIASYLRISEKDEEQIGDDRRVGAIVLAVTILATIIGYAVTNS
 TFPRTIPLQAGLQKPLTPIIEGTAGVGPHVTAELKGGVYKVPGRELTIQVKVTNKTDEPLKL
 GEYTAAGLRFLNPDVFTTKPEFPDYLLADRGLSTDPTPLAPGETKTIEVKVQDARWDIERLSD
 LAYDTSQIGLLMFFSPSGKRYATEIGGPVIPKFVAGDMPGSGQCTNYALLKLAGDVESNP
 GPMASKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWPTL
 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVN
 RIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKANFKTRHNIEDGGVQLADHYQ
 QNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK

Appendix 2: Codon optimized polyMdomain sequence

PolyMdomain:

Pmoc-T2A-pmoAInteinlinker1-P2A-PR1BtruncLinker2Mdomain

>polyMdomain nucleotide sequence

(lower case nucleotides at the beginning and end were added for in frame cloning to N-terminal and C-terminal tags in Gateway destination vectors)

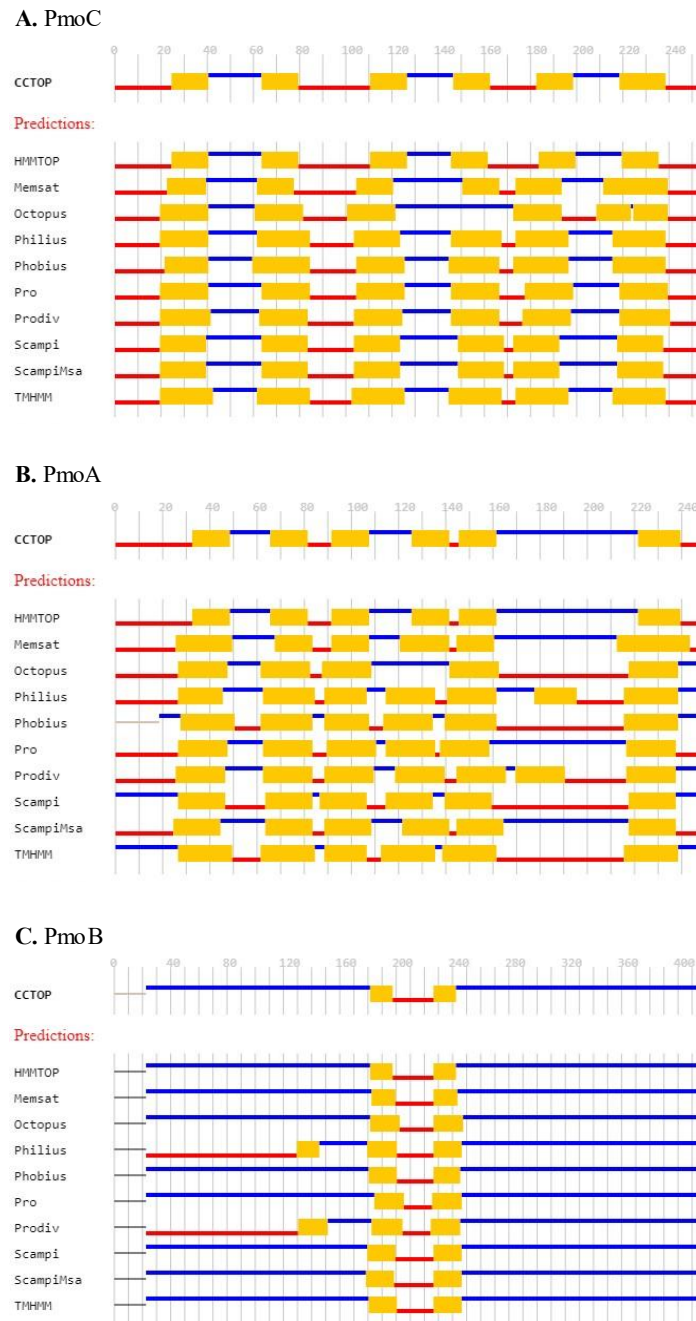
```
ggATGTCTGTCACAACTGAAACAACTGCTGGAGCGGCAGCAGGATCAGACGCGATTGTTG
ACCTTCGGGGCATGTGGGTTGGTGTTCGGGATTGAACATTTTTTATCTGATCGTTCGCAT
ATATGAGCAGATATACGGGTGGAGGGCCGTTTGGATTCTGTTCCGCCCCTGAGTTCCAGAC
ATACTGGCTCAGTATACTCTGGACCGAGATTCCGTTGGAGTTGGTTTCAGGCTTGGCGCT
GGCCGGTTGGTTGTGGAACCAAGAGACAGGAACGTTGACGCCGTGGCACCTAGGGAAG
AGCTCAGGAGACATGTGGTGCTTGTGGAGTGGCTCGTCTGTTACGCCGTGGCTATTTACT
GGGGCGCCTCCTTCTTCACTGAACAAGATGGGACCTGGCACATGACCGTGATACGCGAC
ACAGATTTTACCCCGAGTCATATTATCGAGTTTACATGTCATACCCCATCTATAGTATCA
TGGCGGTGGGAGCGTTCTTCTACGCAAAGACAAGGATCCCGTATTTTCGCGCACGGCTTCT
CCTTGGCCTTCTGATAGTAGCGATCGGCCCTTCATGATCATCCCAAATGTCCGGTCTCA
ATGAATGGGGCCATACGTTTTGTTTTATGGAGGAGCTGTTCTGGCTCCACTCCATTGGG
GGTTCGTCTTCTTCGGCTGGATGGCCCTGGGAGTATTCGGCGTCGTCTGCAGATCCTCAT
GGGTGTCAAGCGTTTGATTGGTAAGGATTGCGTTGCCGCGCTTGTGGGGGAAGTGGCGA
AGGACGCGGTTCTTCTGACGTGTGGTGATGTGGAGGAGAATCCCGGTCCGATGTTTCAC
AAGCAAGAGTGGTGGGGCTATCGGTCCATTTCACTCCGTGCTGAGGCTGCTGGCTGTGT
GAAGACTACGGACTGGATGTTTCTGACACTCCTCTTCTGCGGTATTGGGGGGGTACCA
TATTCATTCATGTTGACGGCCGGAGACTGGGACTTTTGGGTGGATTGGAAGGATAGGCG
TATGTGGCCTACAGTGGTCCCAATCCTGGGTGTGACTTTTGCAGCCGCGGCGCAAGCCTT
TTTCTGGGAGAATTTCAAACCTCCCTTTTCGGAGCAACCTTCGCGGTTTCGGGCCTGCTTATC
GGCGAGTGGATCAACCGCTACTGCAACTTTTGGGGATGGACCTACTTCCCAATTTCACTG
GTGTTCCCGAGCGCTTTGGTGGTTCCAGCACTGTGGCTCGACATAATCATGCTTCTCTCAG
GATCTTACGTAATCACTGCAGTGGTTCGGGTCCCTTGGATGGGGTCTTCTTCTTACCCTAA
TAAGTGGCCAGCTATTGCCGCACTGCACCAAGCGACCGAACAACACGGACAACCTTATGA
GCCTCGCAGACCTCGTGGGGTTTCACTTCGTCCGGACCTCTATGCCGGAATATATCCGCA
TGGTTGAGCGCGGAACATTGCGGACTTTCGGTAAGGAGGTGGTCCCGGTTGCAGCCTTCT
TTTCTGGCTTTGTGTCGATGATGGTATACTTCTCTGTTGGTTCGTGGGGAAATGGTATAG
CACCACCAAAGTGATACAAAAAATATGTTTGTCTTCGGAAGTGAAGATACTTACAGTTGA
ATATGGACCACTTCTATTGGAAAGATTGTGAGTGAAGAGATCAACTGCAGTGTTTATTC
CGTGGATCCAGAGGGTAGAGTTTAACTCAAGCAATTGCTCAGTGGCATGATAGGGGAG
AACAGGAGGTTCTTGAATATGAGTTGGAAGATGGTTCTGTGATAAGAGCTACATCAGAT
CACAGGTTTCTTACTACAGATTACCAACTTTTGGCAATCGAAGAGATTTTCGCTAGACAG
CTCGATCTTCTCACTTTGGAAAATATTAAGCAAACAGAAGAGGCACTTGATAACCATAGG
CTTCCATTTCTCTTTTGGATGCTGGAACATTAAGATGGTTAAAGTGATAGGAAGAAGG
TCATTGGGTGTTCAAAGAATATTTGATATCGGACTTCCTCAGGATCACAAATTTCTTACTCG
CAAACGGTGCTATTGCTGCAGCTGTTCTGTGGTTCTGGTTCTAGAGGTTTCGGGGGCTA
CGAACTTCAGTCTGCTGAAACAGGCGGGAGACGTTGAGGAGAATCCTGCTCCAATGGGA
TTTGTTCCTTTTTTCTCAACTTCCATCTTTTCTTCTTGTTCCTACTCTTCTTCTTTTCT
```

TGTTATTTCTCATTCTTGTAGAGCTCACGGCGAGAAATCGCAGCAGGCGTTCCTCCGCA
TGCGCACTCTCAACTGGTACGATGTAAATGGTCTAAGACATCCCTTAATGTTAATGAAT
CGATGGTCTTTCTGGCAAGGTACACGTCTTTTCGGCGTGGCCCCAGGCAGTTGCGAATC
CTAAAAGCTCCTTCTCAACGCCGGCGAACCTGGCCCTGTTCTTGTTCTGACAGCTCAGTT
CATCGGGGAGCAGTTTGCCCTAGGTACGTGTCATTGGAGGTGGGAAAAGATTACGCGT
TCTCGATAGATCTTAAGGCGCGTAGGGCCGGTAGGTGGCATGTTTCACGCTCAGATTAACG
TAGAGGGCGGAGGTCCTATCATTGGCCCTGGACAGTGGATTGAAATCAAGGGAGACATG
GCAGACTTCAAAGATCCAGTCACATTGCTGGACGGGACAACCGTCGATCTCGAGACTTA
CGGAATCGATCGGATCTACGCCTGGCATTTCCTCATGGATGATAGCTGCTGCCGCGTGGAT
TCTCTATTGGTTCTTTAAAAAGGGGATCATTGCCTCGTACCTGCGCATATCGGAAGGGAA
GGATGAAGAGCAGATTGGTGACGACGACCGGAGGGTCGGTGCAATTGTCCTTGCTGTGA
CAATTTTGGCGACTATCATAGGATACGCTGTGACAAATTCAACCTTCCCAAGGACAATTC
CTCTGCAGGCTGGACTCCAGAAGCCTTTGACGCCGATTATAGAGGAAGGGACGGCGGGA
GTTGGTCCGCATGTGGTTACGGCAGAGCTCAAAGGGGGTGTATACAAAGTCCCGGGCAG
AGAGCTGACGATACAAGTTAAGGTTACCAATAAGACTGATGAGCCCCTCAAGCTCGGAG
AGTATACAGCCGCGGGATTGCGGTTTCTCAATCCTGATGTATTCACGACGAAGCCCGAGT
TTCCAGACTACCTCTTGGCGGACAGAGGACTTTCTACGGATCCTACCCCACTGGCTCCAG
GCGAGACGAAAACGATTGAGGTAAAAGTTCAAGACGCTCGCTGGGATATTGAGCGTCTG
TCGGATCTGGCTTACGACACAGACAGCCAAATAGGCGGCCTCCTGATGTTTTTCTCCCC
TCCGGAAGCGGTACGCCACCGAAATTGGCGGCCCGGTAATTCCGAAATTCGTGGCTGG
AGACATGCCTGGAGGTGGGTCTGGTGGAGGCTCAGCCAACATAACAGTAGACTATCTTT
ACAACAAAGAAACCAAACCTCTTTACTGCTAAGTTGAATGTGAATGAGAATGTTGAATGT
GGAAACAATACATGCACGAATAATGAGGTTCACTTAACCTGAATGCAAGAATGCATC
AGTCAGTATTTCCATAACTCTTGTACTGCTCCAGATc

> PolyMdomain protein sequence:

MSVTTETTAGAAAGSDAIVDLRGMWVGVAAGLNIFYLIVRIYEQIYGWRAGLDSFAPEFQTY
 WLSILWTEIPLLELVSLALAGWLWKTRDRNVDVAPREELRRHVVLVEWLVVYAVAIYWG
 ASFFTEQDGTWHMTVIRDTDFTPSHIIEFYMSYPIYSIMAVGAFFYAKTRIPYFAHGFSLAFLIV
 AIGPFMIIPNVGLNEWGHTFWFMEELFVAPLHWGFVFFGWMALGVFGVVLQILMGVKRLIG
 KDCVAALVGGSGEGRGSLTTCGDVEENPGPMFTSKSGGAIGPFHVSVAEAAAGCVKTTDWMFL
 TLLFLAVLGGYHIHFMLTAGDWDFWVDWKDRRMWPTVVPILGVTFAAAAQAFFWENFKLP
 FGATFAVSGLLIGEWINRYCNFWGWTYFPISLVFPSALVVPALWLDIIMLLSGSYVITAVVGS
 LGWGLLFYPNNWPAIAALHQATEQHGQLMSLADLVGFHFVRTSMPEYIRMVERGTLRTFGK
 EVVPVAAFFSGFVSMMVYFLWWFVGKWYSTTKVIQKICLSFGTEILTVEYGPLPIGKIVSEEI
 NCSVYSVDPEGRVYTQAIQWHDGRGEQEVLEYELEDGSVIRATSDHRFLTDDYQLLAIEEIFA
 RQLDLLTLENIKQTEEALDNHRLPFPLLDAGTIKMKVIGRRSLGVQRIFDIGLPQDHNFLLAN
 GAIAAACSCGSGRGSAGATNFSLLKQAGDVEENPGPMGFVLFSQLPSFLLVSTLLLFLVISH
 SCRAHGEKSQQAFLRMRTLNWYDVKWSKTSLVNENSMVLSGKVHVFSAPWQAVANPKSSF
 LNAGEPGPVLVRTAQFIGEQFAPRSVSLEV GKDYAFSIDLKARRAGRWHVHAQINVEGGGPPII
 GPGQWIEIKGDMADFKDPVTLLDGTTVDLETY GIDRIYAWHFPWMIAAAAWILYWFFKKGII
 ASYLRISEKDEEQIGDDRRVGAIVLAVTILATIIGYAVTNSTFPRTIPLQAGLQKPLTPIIEEG
 TAGVGPHVVTAELKGGVYKVPGRELTIQVKVTNKTDEPLKLGEYTAAGLRFLNPDVFTTKPE
 FPDYLLADRGLSTDPTPLAPGETKTIEVKVQDARWDIERLSLAYDSDS QIGLLMFFSPSGK
 RYATEIGGPVIPKFVAGDMPGGSGGSANITVDYLYNKETKLFTAKLVNENVECGNNTCT
 NNEVHNLTECKNASVSISHNSCTAPD

9.2 Chapter 3 – Expression of pMMO proteins in the plant endoplasmic reticulum (ER)

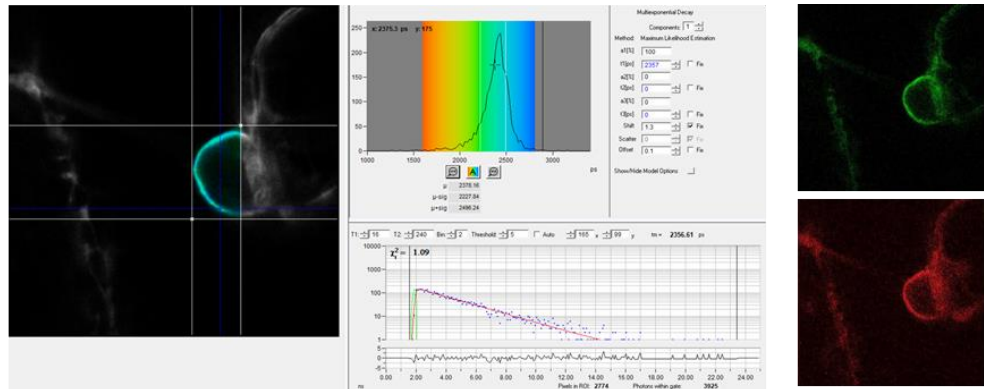


Supplementary Figure 3.1. Output of the CCTOP topology prediction algorithm.

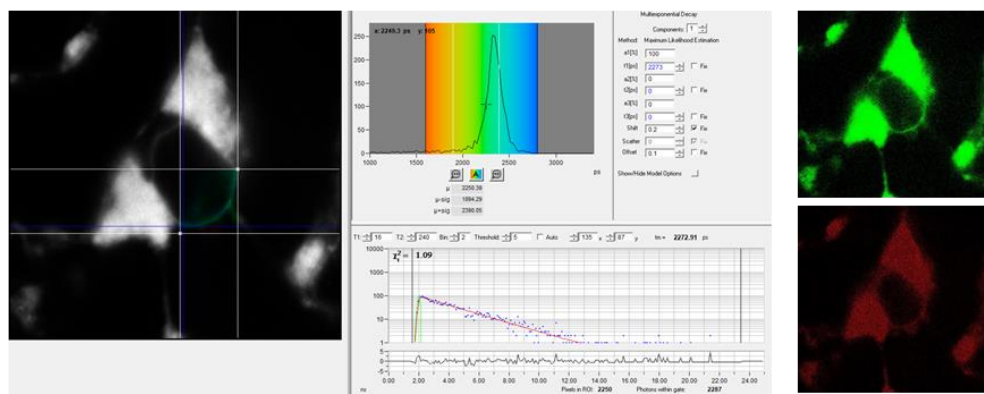
Predictions of topology by 10 different algorithms as well as the CCTOP consensus are shown. Red lines indicate protein regions with a higher likeliness of residing “inside” the cell and blue lines regions residing “outside”. Yellow rectangles represent transmembrane domains. The

grey line at the beginning of spBtrunc indicates the signal peptide. All predictions were scored by the software with high reliability (pmoC: 93.4 %, pmoA: 90.4 % and spBtrunc: 98.9 %).

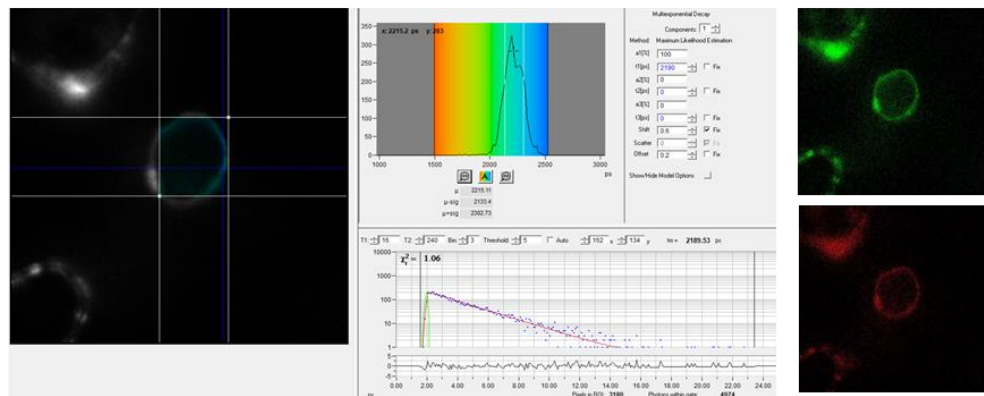
GFP-pmoA +
spBtrunc-RFP Δ del



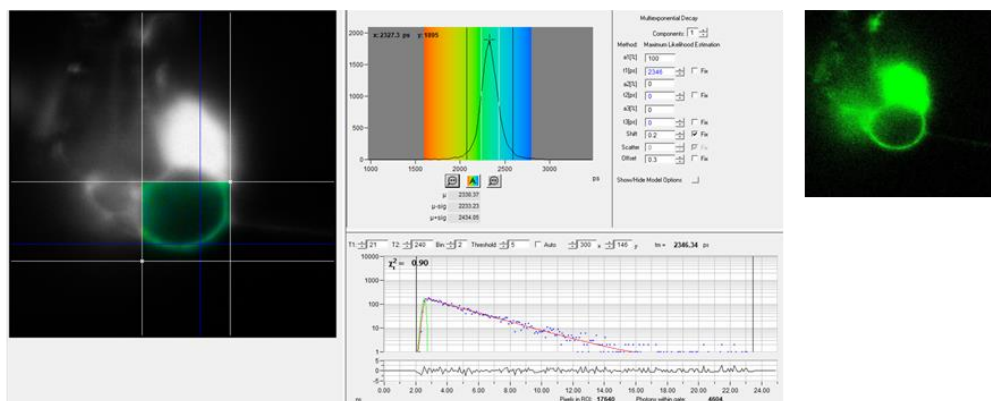
GFP-pmoA +
pmoC-RFP



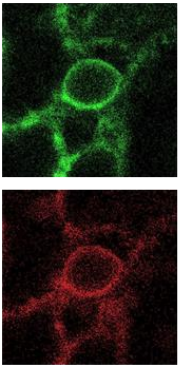
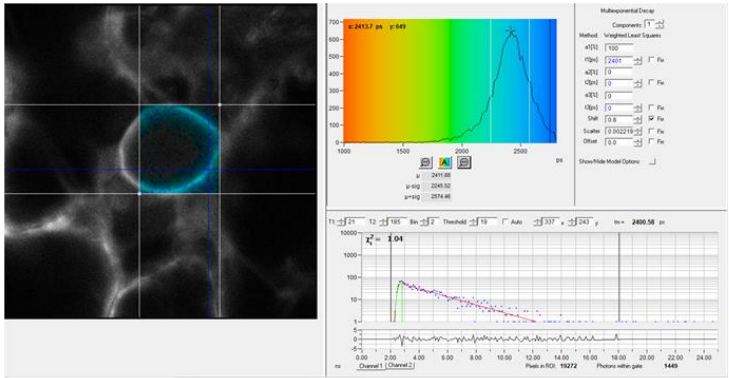
GFP-pmoA +
pmoA-RFP



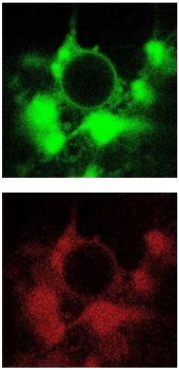
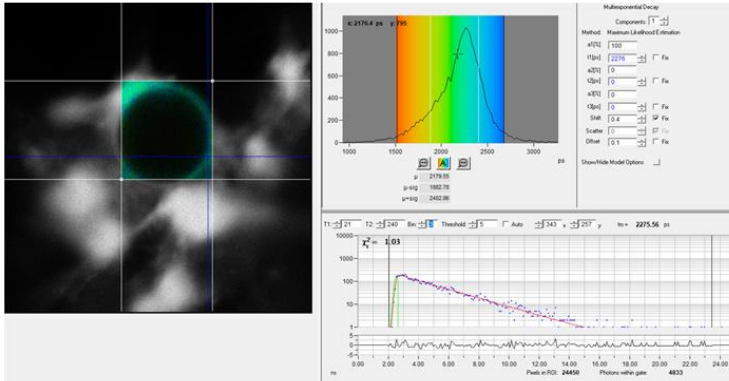
pmoA-GFP
Control



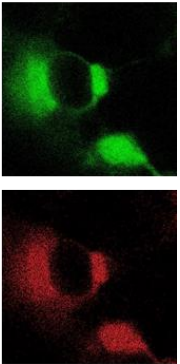
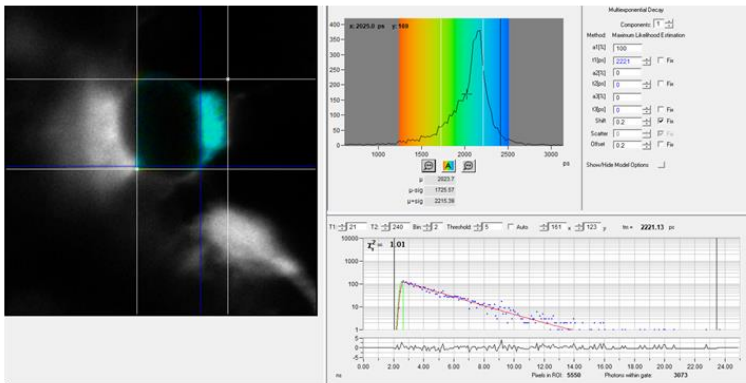
pmoA-GFP +
spBtrunc-RFP-
HDEL



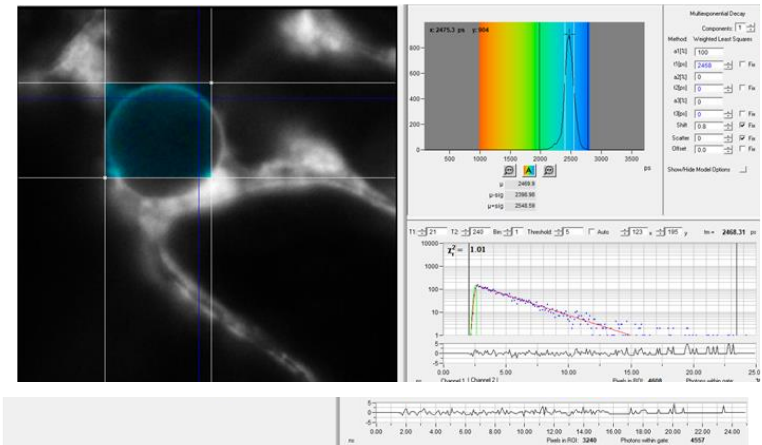
pmoA-GFP +
mCherry-pmoC



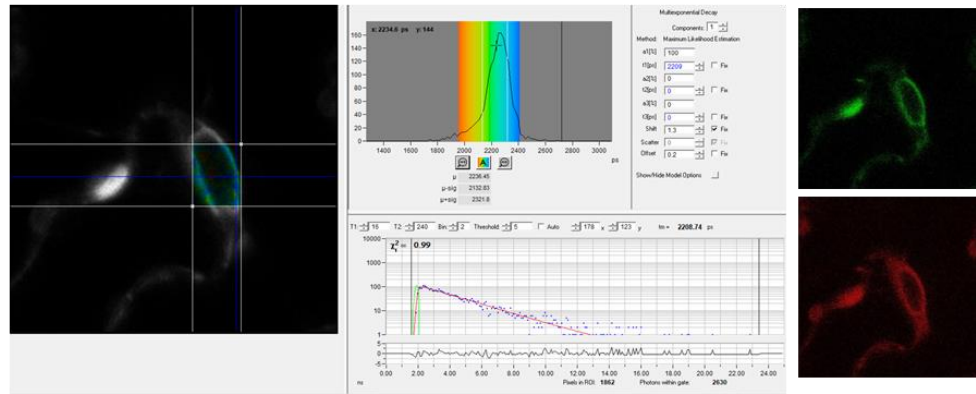
pmoA-GFP +
pmoC-RFP



GFP-pmoC
Control



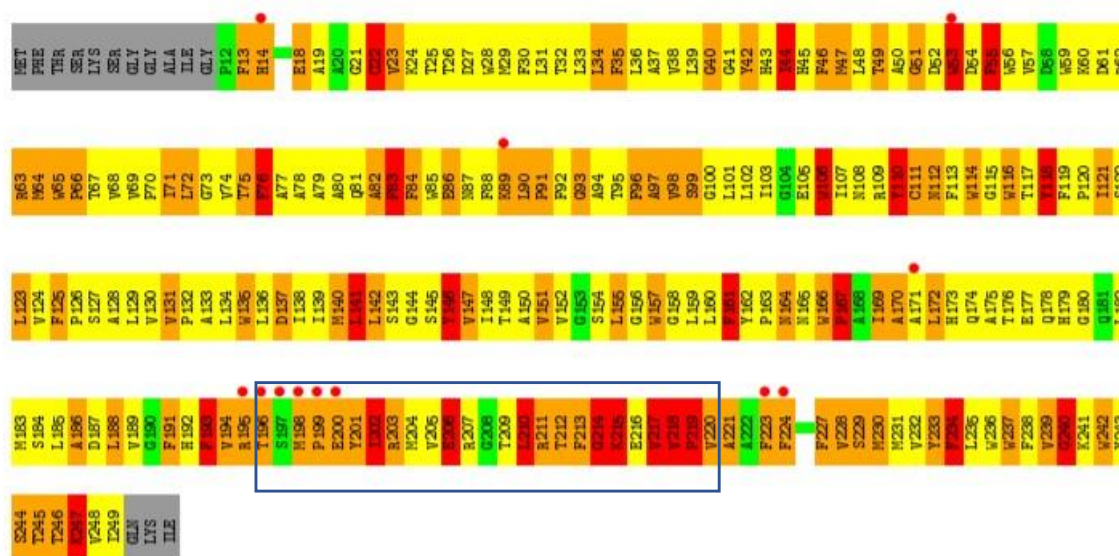
spBtrunc-GFP-
HDEL +
spBtrunc-RFP-
HDEL



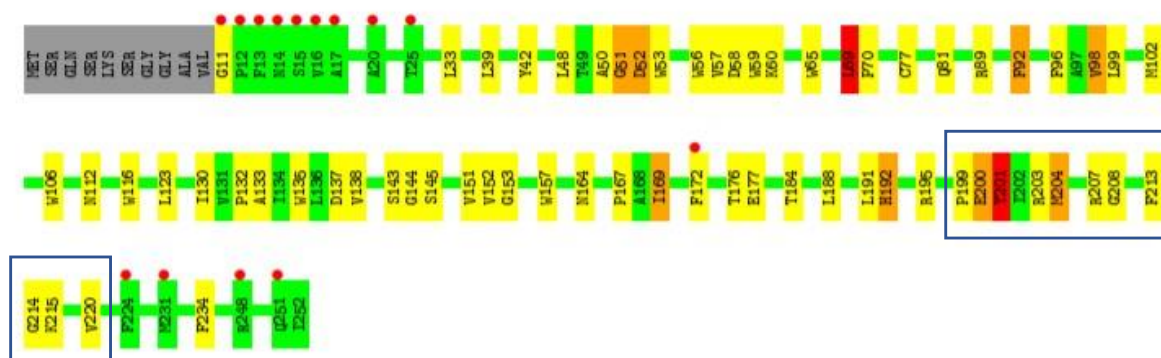
Supplementary Figure 3.2. FRET-FLIM example images and processing.

Example images of each combination used for FRET-FLIM, showing the selected region of the nuclear envelope, the color-coded distribution of the GFP lifetime, the GFP lifetime decay curve for a single pixel within the area and its fit to an exponential curve, and raw images on the right. Images are shown for the donors alone (controls), and in the presence of RFP-tagged acceptors.

A. 3CHX



B. RFR

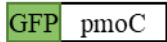
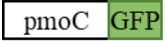
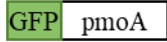
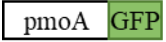
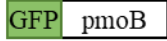
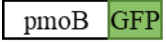
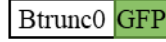
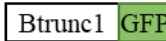
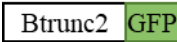


Supplementary Figure 3.3. PmoA structure validation report.

Screenshot of the structure validation report of *pmoA* from the PDB entries 3CHX (A) and 3RFR (B). The color scheme for residues is based on the amount of geometric quality criteria for which they contain outliers (green = 0, yellow = 1, orange = 2 and red = 3 or more). Green lines show 2 or more residues which do not have any outliers. Red dots signify a poor fit to the electron density. The loop region which differs between structure 3CHX and the rest is shown with a rectangle.

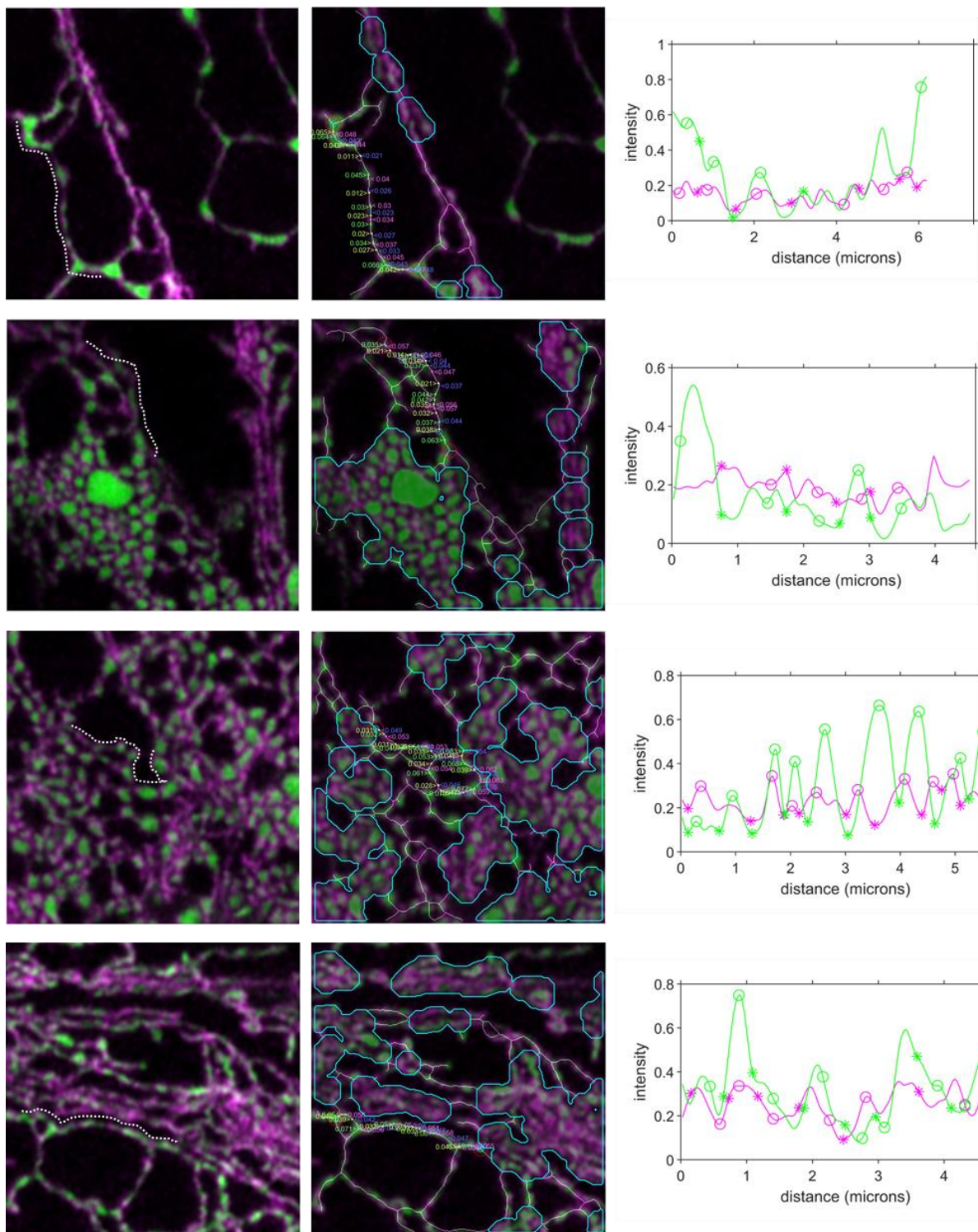
Appendix 3 – Summary of constructs related to targeting and gene stacking from Chapters 3 and 4

Table summarizing constructs created related to subcellular targeting (Chapter 3) and polypeptides (Chapter 4), including their subcellular localization, approximate expression levels without co-expression with P19, and the induced phenotype in the case of severe ER disruption (linked to tissue necrosis at 2 days after infiltration).

Construct	Localisation	Expression levels (without P19)	Induced phenotype
GFP-pmoC 	ER	High	ER disruption
pmoC-GFP 	ER	High	ER disruption
GFP-pmoA 	ER	High	
pmoA-GFP 	ER	High	
GFP-pmoB 	Cytosol	High	
pmoB-GFP 	Cytosol + mitochondria	High	
Btrunc0-GFP 	Cytosol	High	
Btrunc1-GFP (removal of the full bacterial signal peptide) 	Cytosol	Low	
Btrunc2-GFP 	Cytosol	Low	
Btrunc3-GFP	Cytosol	Low	

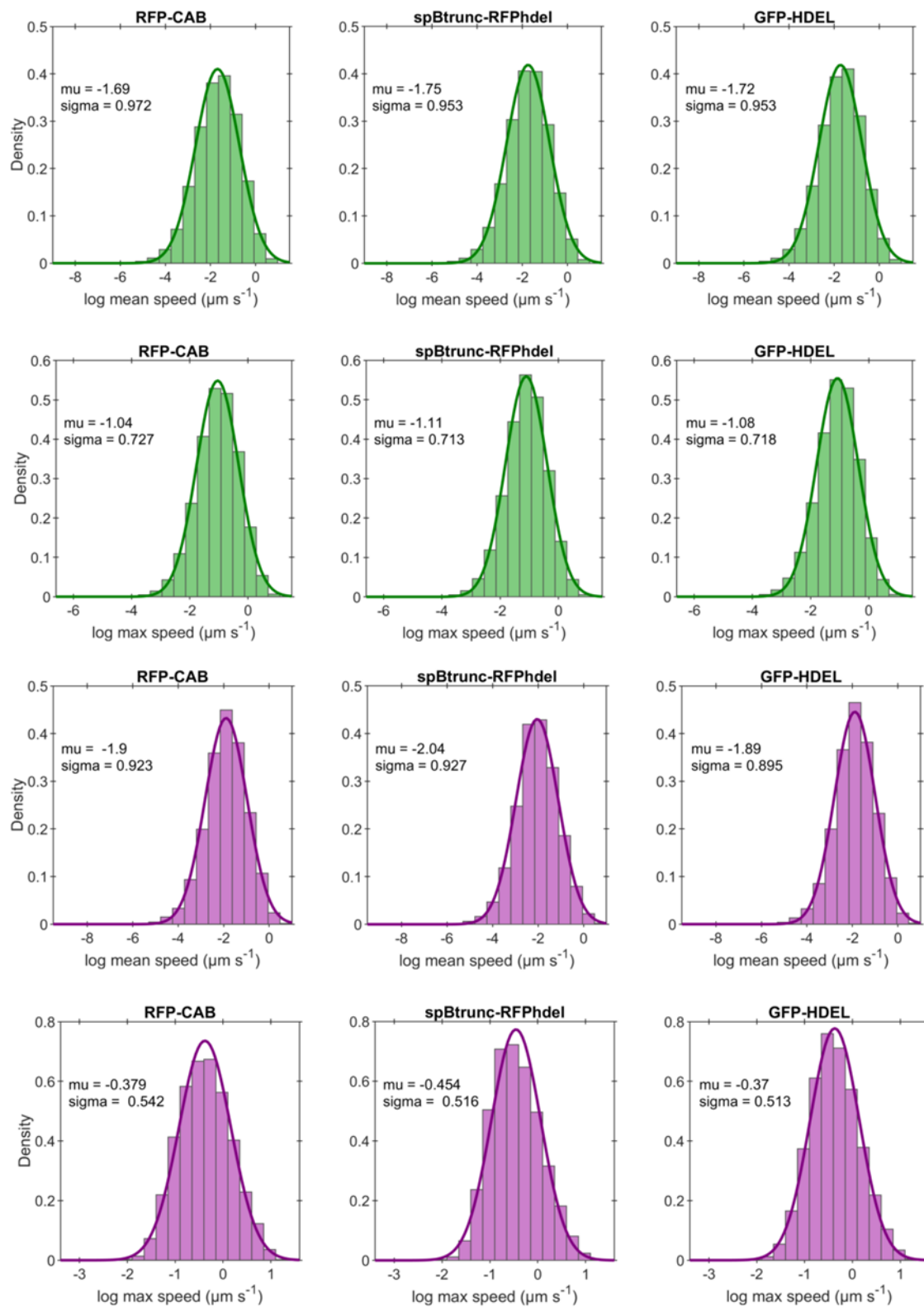
Btrunc3 GFP			
spBtrunc-GFP (plant signal peptide added to Btrunc1) sp Btrunc GFP	ER	Low	
spBtrunc-GFP ~~hdel~~ sp Btrunc GFP ~~hdel~~	ER	Low	
spBtrunc-GFP ~~hdel2~~ sp Btrunc GFP ~~hdel2~~	ER	Low	
spBtrunc-RFP sp Btrunc RFP	ER + vacuole	High	
spBtrunc-RFP ~~hdel~~ sp Btrunc RFP ~~hdel~~	ER	High	
spBtrunc-RFP-KKRY sp Btrunc RFP kkry	ER + vacuole	High	
RFP-polypeptide RFP pmoC T2A pmoA P2A spBtrunc E2A Clover	ER(pmoC)/ cytosol (clover)	Low	
RFP-CAB RFP pmoC T2A pmoA P2A spBtrunc	ER	Low	
RFP-CA RFP pmoC T2A pmoA	ER	High	ER disruption
RFP-CAB-GFP RFP pmoC T2A pmoA P2A spBtrunc GFP	ER(pmoC)/ ER(spBtrunc)	Low	
RFP-CA-GFP RFP pmoC T2A pmoA GFP	ER(pmoC)/ ER(pmoA)	High	ER disruption

9.3 Chapter 5 – Impact of pMMO on ER structure and dynamics



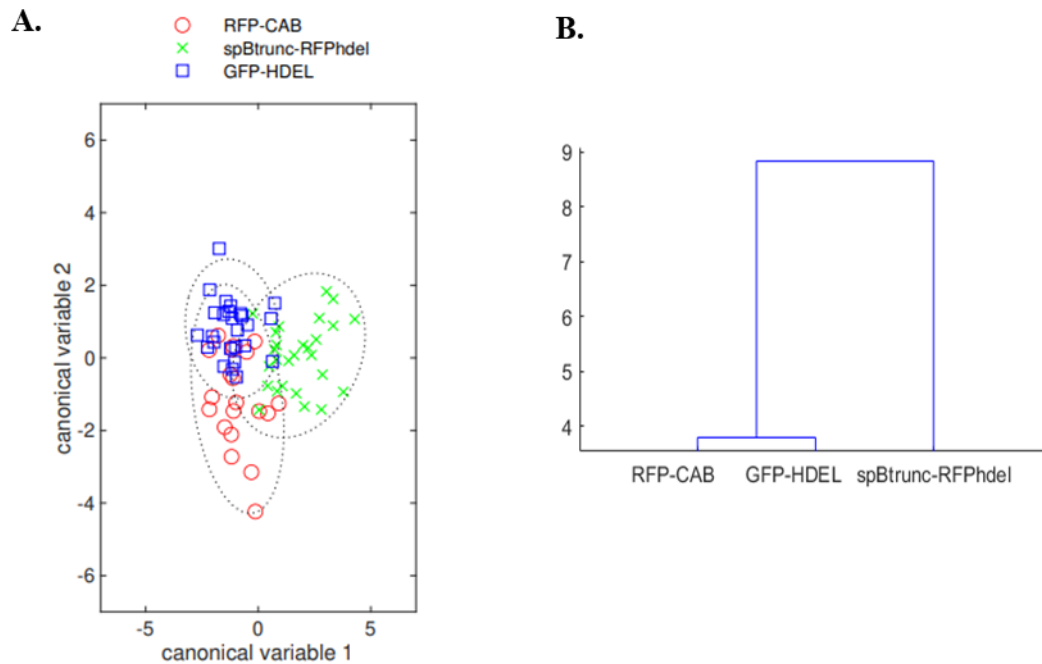
Supplementary Figure 5.1 Tubule traces.

High magnification images of four independent cells expressing RFP-CAB + GFP-HDEL showing a patchy phenotype of the luminal marker. Trace plots along a tubule (marked with a white dotted line) show the normalized intensity values in the red and green channels, circles and stars mark the peaks and troughs. On some occasions a peak in the green channel coincides with a trough in the red channel, however, in other cases the intensity in the red remains constant during a peak in the green, and in other cases both peaks coincide, differing from a constriction phenotype in which both intensities show a clear negative correlation.



Supplementary Figure 5.2. Log transformed distributions of speed.

Log normal maximum and mean speeds in tubules (green) and cisternae (magenta) showing the mean and standard deviation, for all three groups. The mean values were back-transformed to obtain the values of speeds.



Supplementary Figure 5.3. Multivariate analysis using distance = 3 for cisternal correlation.

(A) MANOVA using a subset of 19 metrics for ER structure and dynamics, showing a separation between all three groups in 2 canonical variables (Pillai's trace, $F(3.0155)$, $p = 5.7 \times 10^{-6}$; Roy's largest root, $F(1.9784)$, $p = 1.7 \times 10^{-7}$). Dotted lines represent the 95 % confidence limit. (B) Dendrogram plotting the similarity between groups based on the multivariate means.

References

Almagro Armenteros JJ, Sønderby CK, Sønderby SK, Nielsen H, Winther O. 2017. DeepLoc: prediction of protein subcellular localization using deep learning. *Bioinformatics* (Oxford, England) **33**, 3387–3395.

Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology* **37**, 420–423.

El Amrani A, Barakate A, Askari BM, Li X, Roberts AG, Ryan MD, Halpin C. 2004. Coordinate expression and independent subcellular targeting of multiple proteins from a single transgene. *Plant physiology* **135**, 16–24.

Annual Review: MX Diamond Light Source.

Arongaus AB, Chen S, Pireyre M, Glöckner N, Galvão VC, Albert A, Winkler JB, Fankhauser C, Harter K, Ulm R. 2018. Arabidopsis RUP2 represses UVR8-mediated flowering in noninductive photoperiods. *Genes & development* **32**, 1332–1343.

Baani M, Liesack W. 2008. Two isozymes of particulate methane monooxygenase with different methane oxidation kinetics are found in *Methylocystis* sp. strain SC2. *Proceedings of the National Academy of Sciences* **105**, 10203–10208.

Bai L, You Q, Feng X, Kovach A, Li H. 2020. Structure of the ER membrane complex, a transmembrane-domain insertase. *Nature* 2020 584:7821 **584**, 475–478.

Bajar BT, Wang ES, Zhang S, Lin MZ, Chu J. 2016. A guide to fluorescent protein FRET pairs. *Sensors* (Basel, Switzerland) **16**.

Balasubramanian R, Smith SM, Rawat S, Yatsunyk LA, Stemmler TL, Rosenzweig AC. 2010. Oxidation of methane by a biological dicopper centre. *Nature* **465**, 115–119.

Barbante A, Irons S, Hawes C, Frigerio L, Vitale A, Pedrizzini E. 2008. Anchorage to the cytosolic face of the endoplasmic reticulum membrane: a new strategy to stabilize a cytosolic recombinant antigen in plants. *Plant Biotechnology Journal* **6**, 560–575.

Batoko H, Zheng HQ, Hawes C, Moore I. 2000. A Rab1 GTPase is required for transport between the endoplasmic reticulum and golgi apparatus and for normal Golgi movement in plants. *Plant Cell* **12**, 2201–2217.

Benghezal M, Wasteneys GO, Jones DA. 2000. The C-terminal dilysine motif confers

endoplasmic reticulum localization to type I membrane proteins in plants. *Plant Cell* **12**, 1179–1201.

Bharali A, Baruah KK, Gogoi N. 2017. Potential option for mitigating methane emission from tropical paddy rice through selection of suitable rice varieties. *Crop and Pasture Science* **68**, 421.

Bhattacharyya P, Dash PK, Swain CK, et al. 2018. Mechanism of plant mediated methane emission in tropical lowland rice. *Science of the Total Environment* **651**, 84–92.

Bhoo SH, Lai H, Ma J, Arntzen CJ, Chen Q, Mason HS. 2011. Expression of an immunogenic Ebola immune complex in *Nicotiana benthamiana*. *Plant Biotechnology Journal* **9**, 807.

Bhullar GS, Edwards PJ, Olde Venterink H. 2013. Variation in the plant-mediated methane transport and its importance for methane emission from intact wetland peat mesocosms. *Journal of Plant Ecology* **6**, 298–304.

Biswal AK, Shamim M, Cruzado K, Soriano G, Ghatak A, Toleco MR, Vikram P. 2017. Role of biotechnology in rice production. *Rice Production Worldwide*, 487–547.

Bohmert-Tatarev K, McAvoy S, Daughtry S, Peoples OP, Snell KD. 2011. Focus issue on plastid biology: high levels of bioplastic are produced in fertile transplastomic tobacco plants engineered with a synthetic operon for the production of polyhydroxybutyrate. *Plant Physiology* **155**, 1690.

Borgese N, Francolini M, Snapp E. 2006. Endoplasmic reticulum architecture: structures in flux. *Current Opinion in Cell Biology* **18**, 358.

Botchway SW, Scherer KM, Hook S, Stubbs CD, Weston E, Bisby RH, Parker AW. 2015. A series of flexible design adaptations to the Nikon E-C1 and E-C2 confocal microscope systems for UV, multiphoton and FLIM imaging. *Journal of Microscopy* **258**, 68–78.

Boulafloous A, Saint-Jore-Dupas C, Herranz-Gordo MC, Pagny-Salehabadi S, Plasson C, Garidou F, Kiefer-Meyer MC, Ritzenthaler C, Faye L, Gomord V. 2009. Cytosolic N-terminal arginine-based signals together with a luminal signal target a type II membrane protein to the plant ER. *BMC Plant Biology* **9**, 144.

Braakman I, Hebert DN. 2013. Protein folding in the endoplasmic reticulum. *Cold Spring Harbor Perspectives in Biology* **5**.

Brach T, Soyk S, Müller C, Hinz G, Hell R, Brandizzi F, Meyer AJ. 2009. Non-invasive topology analysis of membrane proteins in the secretory pathway. *The Plant Journal* **57**, 534–541.

Brandizzi F. 2021. Maintaining the structural and functional homeostasis of the plant endoplasmic reticulum. *Developmental Cell* **56**, 919–932.

Brandizzi F, Hanton S, DaSilva LLP, Boevink P, Evans D, Oparka K, Denecke J, Hawes C. 2003. ER quality control can lead to retrograde transport from the ER lumen to the cytosol and the nucleoplasm in plants. *The Plant Journal* **34**, 269–281.

Breeze E, Dzimitrowicz N, Kriechbaumer V, et al. 2016. A C-terminal amphipathic helix is necessary for the *in vivo* tubule-shaping function of a plant reticulon. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 10902–10907.

Broussard JA, Green KJ. 2017. Research techniques made simple: methodology and applications of Förster Resonance Energy Transfer (FRET) microscopy. *The Journal of Investigative Dermatology* **137**, e185–e191.

Burén S, Ortega-Villasante C, Ötvös K, Samuelsson G, Bakó L, Villarejo A. 2012. Use of the Foot-and-Mouth Disease virus 2A peptide co-expression system to study intracellular protein trafficking in *Arabidopsis* (S Caplan, Ed.). *PLoS ONE* **7**, e51973.

Butaye KMJ, Goderis IJWM, Wouters PFJ, Poes JMTG, Delauré SL, Broekaert WF, Depicker A, Cammue BPA, De Bolle MFC. 2004. Stable high-level transgene expression in *Arabidopsis thaliana* using gene silencing mutants and matrix attachment regions. *The Plant Journal* **39**, 440–449.

Butterbach-Bahl K, Papen H, Rennenberg H. 1997. Impact of gas transport through rice cultivars on methane emission from rice paddy fields. *Plant, Cell and Environment* **20**, 1175–1183.

Candat A, Paszkiewicz G, Neveu M, Gautier R, Logan DC, Avelange-Macherel M-H, Macherel D. 2014. The ubiquitous distribution of late embryogenesis abundant proteins across cell compartments in *Arabidopsis* offers tailored protection against abiotic stress. *The Plant Cell* **26**, 3148–3166.

Cao L, Caldararu O, Rosenzweig AC, Ryde U. 2018. Quantum refinement does not support dinuclear copper sites in crystal structures of particulate methane monooxygenase. *Angewandte Chemie* **130**, 168–172.

Cao J, Zhao JZ, Tang JD, Shelton AM, Earle ED. 2002. Broccoli plants with pyramided cry1Ac and cry1C Bt genes control diamondback moths resistant to Cry1A and Cry1C proteins. *Theoretical and Applied Genetics* **105**, 258–264.

Ceccon CC, Caverzan A, Margis R, Salvadori JR, Grandó MF. 2020. Gene stacking as a strategy to confer characteristics of agronomic importance in plants by genetic engineering. *Ciência Rural* **50**.

Chan SI, Chang WH, Huang SH, Lin HH, Yu SSF. 2021. Catalytic machinery of methane oxidation in particulate methane monooxygenase (pMMO). *Journal of Inorganic Biochemistry* **225**, 111602.

Chandra R, Takeuchi H, Hasegawa T. 2012. Methane production from lignocellulosic agricultural crop wastes: A review in context to second generation of biofuel production. *Renewable and Sustainable Energy Reviews* **16**, 1462–1476.

Chang W-H, Lin H-H, Tsai I-K, Huang S-H, Chung S-C, Tu I-P, Yu S-F, Chan SI. 2021. Copper centers in the cryo-EM structure of particulate methane monooxygenase reveal the catalytic machinery of methane oxidation. *J. Am. Chem. Soc* **143**, 59.

Chen S, Desai T, McNew JA, Gerard P, Novick PJ, Ferro-Novick S. 2015. Lunapark stabilizes nascent three-way junctions in the endoplasmic reticulum. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 418–423.

Chen S, Novick P, Ferro-Novick S. 2012. ER network formation requires a balance of the dynamin-like GTPase Sey1p and Lnp1p, a member of the Lunapark family. *Nature Cell Biology* **14**, 707.

Chen Y, Zhang Y, Li S, et al. 2021. OsRGA1 optimizes photosynthate allocation for roots to reduce methane emissions and improve yield in paddy ecosystems. *Soil Biology and Biochemistry* **160**, 108344.

Cherkas V, Grebenyuk S, Osypenko D, et al. 2018. Measurement of intracellular concentration of fluorescently-labeled targets in living cells. *PLOS ONE* **13**, e0194031.

Choi DW, Kunz RC, Boyd ES, Semrau JD, Antholine WE, Han JI, Zahn JA, Boyd JM, De la Mora AM, DiSpirito AA. 2003. The membrane-associated methane monooxygenase (pMMO) and pMMO-NADH:Quinone oxidoreductase complex from *Methylococcus capsulatus* Bath. *Journal of Bacteriology* **185**, 5755–5764.

Clough SJ, Bent AF. 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* **16**, 735–743.

Conrad U, Fiedler U. 1998. Compartment-specific accumulation of recombinant immunoglobulins in plant cells: an essential tool for antibody production and immunomodulation of physiological functions and pathogen activity. *Plant Molecular Biology* **38**, 101–109.

Corry B, Jayatilaka D, Martinac B, Rigby P. 2006. Determination of the orientational distribution and orientation factor for transfer between membrane-bound fluorophores using a confocal microscope. *Biophysical Journal* **91**, 1032–1045.

De Cosa B, Moar W, Lee SB, Miller M, Daniell H. 2001. Overexpression of the Bt cry2Aa2

operon in chloroplasts leads to formation of insecticidal crystals. *Nature Biotechnology* **19**, 71–74.

Cossins EA. 1964. The utilization of carbon-14 compounds by plants: I. The metabolism of methanol-C¹⁴ and its role in amino acid biosynthesis. *Canadian Journal of Biochemistry* **42**, 1793–1802.

Cutsail GE, Ross MO, Rosenzweig AC, Debeer S. 2021. Towards a unified understanding of the copper sites in particulate methane monooxygenase: an X-ray absorption spectroscopic investigation. *Chemical Science* **12**, 6194–6209.

Dastmalchi M, Bernards MA, Dhaubhadel S. 2016. Twin anchors of the soybean isoflavonoid metabolon: evidence for tethering of the complex to the endoplasmic reticulum by IFS and C4H. *The Plant Journal* **85**, 689–706.

Datta R, Heaster TM, Sharick JT, Gillette AA, Skala MC. 2020. Fluorescence lifetime imaging microscopy: fundamentals and advances in instrumentation, analysis, and applications. *Journal of Biomedical Optics* **25**, 071203.

Denecke J, De Rycke R, Botterman J. 1992. Plant and mammalian sorting signals for protein retention in the endoplasmic reticulum contain a conserved epitope. *EMBO Journal* **11**, 2345–2355.

Diamond Light Source. 2015. <https://www.diamond.ac.uk/Home/Corporate-Literature/Annual-Review/Review2016/Villages/Macromolecular-Crystallography-Village/A-new-family-of-copper-storage-proteins-discovered-in-methaneoxidising-bacteria.html>

Dickey A, Wang N, Cooper E, Tull L, Breedlove D, Mason H, Liu D, Wang KY. 2017. Transient expression of lumbrokinase (PI239) in tobacco (*Nicotiana tabacum*) using a geminivirus-based single replicon system dissolves fibrin and blood clots. *Evidence-based Complementary and Alternative Medicine* **2017**.

Dlugokencky EJ. 2021. Global Monitoring Laboratory - Carbon Cycle Greenhouse Gases. US Department of Commerce, NOAA.

Dlugokencky EJ, Nisbet EG, Fisher R, Lowry D. 2011. Global atmospheric methane: budget, changes and dangers. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **369**, 2058–2072.

Donnelly MLL, Luke G, Mehrotra A, Li X, Hughes LE, Gani D, Ryan MD. 2001. Analysis of the aphthovirus 2A/2B polyprotein ‘cleavage’ mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal ‘skip’. *Journal of General Virology* **82**, 1013–1025.

Du L, Wang Y, Shan Z, Shen X, Wang F, Su J. 2021. Comprehensive analysis of SUSIBA2 rice: The low-methane trait and associated changes in soil carbon and microbial communities. *Science of the Total Environment* **764**.

E. Bjorck C, D. Dobson P, Pandhal J. 2018. Biotechnological conversion of methane to methanol: evaluation of progress and potential. *AIMS Bioengineering* **5**, 1–38.

Engler C, Kandzia R, Marillonnet S. 2008. A one pot, one step, precision cloning method with high throughput capability. *PLoS ONE* **3**, 3647.

Farnebäck G. 2003. Two-frame motion estimation based on polynomial expansion. *Image Analysis: Proceedings of the 13th Scandinavian Conference, SCIA*, 363–370.

Fiedler U, Phillips J, Artsaenko O, Conrad U. 1997. Optimization of scFv antibody production in transgenic plants. *Immunotechnology : an international journal of immunological engineering* **3**, 205–216.

Fricker MD. 2016. Quantitative redox imaging software. *Antioxidants and Redox Signaling* **24**, 752–762.

Friedman JR, Voeltz GK. 2011. The ER in 3-D: a multifunctional dynamic membrane network. *Trends in Cell Biology* **21**, 709.

Gaber A, Kim SJ, Kaake RM, Benčina M, Krogan N, Šali A, Pavšič M, Lenarčič B. 2018. EpCAM homo-oligomerization is not the basis for its role in cell-cell adhesion. *Scientific Reports* 2018 8:1 **8**, 1–16.

Gao SJ, Damaj MB, Park JW, et al. 2013. Enhanced transgene expression in sugarcane by co-expression of virus-encoded RNA silencing suppressors. *PLOS ONE* **8**, e66046.

Gnanasekaran T, Karcher D, Nielsen AZ, et al. 2016. Transfer of the cytochrome P450-dependent dhurrin pathway from *Sorghum bicolor* into *Nicotiana tabacum* chloroplasts for light-driven synthesis. *Journal of Experimental Botany* **67**, 2495–2506.

Gottschamel J, Waheed MT, Clarke JL, Lössl AG. 2013. A novel chloroplast transformation vector compatible with the Gateway® recombination cloning technology. *Transgenic Research* **22**, 1273–1278.

Gou Z, Xing XH, Luo M, Jiang H, Han B, Wu H, Wang L, Zhang F. 2006. Functional expression of the particulate methane mono-oxygenase gene in recombinant *Rhodococcus erythropolis*. *FEMS Microbiology Letters* **263**, 136–141.

Gouiaa S, Khoudi H, Leidi EO, Pardo JM, Masmoudi K. 2012. Expression of wheat Na⁺/H⁺ antiporter TNHXS1 and H⁺ pyrophosphatase TVP1 genes in tobacco from a bicistronic

transcriptional unit improves salt tolerance. *Plant Molecular Biology* **79**, 137–155.

Gout E, Aubert S, Bligny R, Rébeillé F, Nonomura AR, Benson AA, Douce R. 2000. Metabolism of methanol in plant cells. Carbon-13 nuclear magnetic resonance studies. *Plant Physiology* **123**, 287–296.

Greenhouse Procedures Manual | McCouch RiceLab.
http://ricelab.plbr.cornell.edu/greenhouse_procedures_manual

Grefen C, Blatt M. 2012. A 2in1 cloning system enables ratiometric bimolecular fluorescence complementation (rBiFC). *BioTechniques* **53**.

Griffing LR, Lin C, Perico C, White RR, Sparkes I. 2017. Plant ER geometry and dynamics: biophysical and cytoskeletal control during growth and biotic response. *Protoplasma* **254**, 43–56.

Grisshammer R. 2006. Understanding recombinant expression of membrane proteins. *Current Opinion in Biotechnology* **17**, 337–340.

Van Groenigen KJ, Van Kessel C, Hungate BA. 2012. Increased greenhouse-gas intensity of rice production under future atmospheric conditions. *Nature Climate Change* 2012 3:3 **3**, 288–291.

Gutsche N, Holtmannspötter M, Maß L, O'Donoghue M, Busch A, Lauri A, Schubert V, Zachgo S. 2017. Conserved redox-dependent DNA binding of ROXY glutaredoxins with TGA transcription factors. *Plant Direct* **1**, e00030.

Ha S-H, Liang YS, Jung H, Ahn M-J, Suh S-C, Kweon S-J, Kim D-H, Kim Y-M, Kim J-K. 2010. Application of two bicistronic systems involving 2A and IRES sequences to the biosynthesis of carotenoids in rice endosperm. *Plant Biotechnology Journal* **8**, 928–938.

Hakemian AS, Kondapalli KC, Telser J, Hoffman BM, Stemmler TL, Rosenzweig AC. 2008. The metal centers of particulate methane monooxygenase from *Methylosinus trichosporium OB3b*. *Biochemistry* **11**, 58.

Halpin C. 2005. Gene stacking in transgenic plants – the challenge for 21st century plant biotechnology. *Plant Biotechnology Journal* **3**, 141–155.

Han X, Sun X, Wang C, Wu M, Dong D, Zhong T, Thies JE, Wu W. 2016. Mitigating methane emission from paddy soil with rice-straw biochar amendment under projected climate change. *Scientific Reports* **6**, 1–10.

Hanson GT, Aggeler R, Oglesbee D, Cannon M, Capaldi RA, Tsien RY, Remington SJ. 2004. Investigating mitochondrial redox potential with redox-sensitive green fluorescent

protein indicators. *Journal of Biological Chemistry* **279**, 13044–13053.

Hara-Hishimura I, Takeuchi Y, Inoue K, Nishimura M. 1993. Vesicle transport and processing of the precursor to 2S albumin in pumpkin. *Plant Journal* **4**, 793–800.

Hecker A, Wallmeroth N, Peter S, Blatt MR, Harter K, Grefen C. 2015. Binary 2in1 vectors improve in planta (co)localization and dynamic protein interaction studies. *Plant Physiology* **168**, 776–787.

Hee Kim J, Lee S-R, Li L-H, Park H-J, Park J-H, Youl Lee K, Kim M-K, Ahn Shin B, Choi S-Y. 2011. High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* **6**.

Hegde RS, Keenan RJ. 2021. The mechanisms of integral membrane protein biogenesis. *Nature Reviews Molecular Cell Biology*. 1–18.

von Heijne G. 1985. Signal sequences. The limits of variation. *Journal of Molecular Biology* **184**, 99–105.

Houbaert A, Zhang C, Tiwari M, Wang K, Marcos Serrano A, Savatin DV, Urs MJ, Zhiponova MK, Gudesblat GE, Vanhoutte I, Eechhout D, Boeren S, Karimin M, Betti C, Jacobs T, Fenoll C, Mena M, de Vries S, De Jaeger G, Russinova E. 2018. POLAR-guided signalling complex assembly and localization drive asymmetric cell division. *Nature* **563**, 574–578.

Huber-Humer M, Gebert J, Hilger H. 2008. Biotic systems to mitigate landfill methane emissions. *Waste Management and Research* **26**, 33–46.

Huff J. 2015. The Airyscan detector from ZEISS: confocal imaging with improved signal-to-noise ratio and super-resolution. *Nature Methods* 2015 12:12 **12**, i–ii.

Hussain S, Peng S, Fahad S, Khaliq A, Huang J, Cui K, Nie L. 2015. Rice management interventions to mitigate greenhouse gas emissions: a review. *Environmental Science and Pollution Research* **22**, 3342–3360.

Irons SL, Evans DE, Brandizzi F. 2003. The first 238 amino acids of the human lamin B receptor are targeted to the nuclear envelope in plants. *Journal of Experimental Botany* **54**, 943–950.

Jackson RB, Sauniois M, Bousquet P, Canadell JG, Poulter B, Stavert AR, Bergamaschi P, Niwa Y, Segers A, Tsuruta A. 2020. Increasing anthropogenic methane emissions arise equally from agricultural and fossil fuel sources. *Environmental Research Letters* **15**, 071002.

Jarvis P. 2008. Targeting of nucleus-encoded proteins to chloroplasts in plants. *New*

Phytologist **179**, 257–285.

Jodts RJ, Ross MO, Koo CW, Doan PE, Rosenzweig AC, Hoffman BM. 2021. Coordination of the copper centers in particulate methane monooxygenase: comparison between methanotrophs and characterization of the CuCSite by EPR and ENDOR spectroscopies. *Journal of the American Chemical Society* **143**, 15358–15368.

Ju C, Chang C. 2012. Advances in ethylene signalling: protein complexes at the endoplasmic reticulum membrane. *AoB PLANTS* **2012**, pls031–pls031.

Kalyuzhnaya MG, Puri AW, Lidstrom ME. 2015. Metabolic engineering in methanotrophic bacteria. *Metabolic Engineering* **29**, 142–152.

Kang H, Park Y, Lee Y, Yoo YJ, Hwang I. 2018. Fusion of a highly N-glycosylated polypeptide increases the expression of ER-localized proteins in plants. *Scientific Reports* **8**, 1–10.

Karimi M, Inzé D, Depicker A. 2002. GATEWAYTM vectors for Agrobacterium-mediated plant transformation. *Trends in Plant Science* **7**, 193–195.

Kavitha B, Reddy PVL, Kim B, Lee SS, Pandey SK, Kim KH. 2018. Benefits and limitations of biochar amendment in agricultural soils: A review. *Journal of Environmental Management* **227**, 146–154.

Kerselidou D, Saeed Dohai B, Nelson DR, et al. 2020. Exostosin-1 glycosyltransferase regulates endoplasmic reticulum architecture and dynamics. *bioRxiv*, 2020.09.02.275925.

Kitmitto A, Myronova N, Basu P, Dalton H. 2005. Characterization and structural analysis of an active particulate methane monooxygenase trimer from *Methylococcus capsulatus* (Bath). *Biochemistry* **44**, 10954–10965.

Koelbener A, Ström L, Edwards PJ, Olde Venterink H. 2010. Plant species from mesotrophic wetlands cause relatively high methane emissions from peat soil. *Plant and Soil* **326**, 147–158.

Koo CW, Rosenzweig AC. 2021. Biochemistry of aerobic biological methane oxidation. *Chemical Society Reviews* **50**, 3424–3436.

Kriechbaumer V, Botchway SW, Hawes C. 2016. Localization and interactions between Arabidopsis auxin biosynthetic enzymes in the TAA/YUC-dependent pathway. *Journal of Experimental Botany* **67**, 4195–4207.

Kriechbaumer V, Botchway SW, Slade SE, Knox K, Frigerio L, Oparka K, Hawes C. 2015. Reticulomics: Protein-protein interaction studies with two plasmodesmata-localized

reticulon family proteins identify binding partners enriched at plasmodesmata, endoplasmic reticulum, and the plasma membrane 1. *Plant Physiology* **169**, 1933–1945.

Kriechbaumer V, Brandizzi F. 2020. The plant endoplasmic reticulum: an organized chaos of tubules and sheets with multiple functions. *Journal of Microscopy* **280**, 122–133.

Kriechbaumer V, Breeze E, Pain C, Tolmie F, Frigerio L, Hawes C. 2018. Arabidopsis Lunapark proteins are involved in ER cisternae formation. *New Phytologist* **219**, 990–1004.

Kritee K, Nair D, Zavala-Araiza D, et al. 2018. High nitrous oxide fluxes from rice indicate the need to manage water for both long- and short-term climate impacts. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 9720–9725.

Lawton TJ, Rosenzweig AC. 2016. Methane-oxidizing enzymes: An upstream problem in biological gas-to-liquids conversion. *Journal of the American Chemical Society* **138**, 9327–40.

Lazareva EA, Lezzhov AA, Chergintsev DA, Golyshev SA, Dolja V V., Morozov SY, Heinlein M, Solovyev AG. 2021. Reticulon-like properties of a plant virus-encoded movement protein. *New Phytologist* **229**, 1052–1066.

Lee J, Freddolino PL, Zhang Y. 2017. Ab initio protein structure prediction. From protein structure to function with bioinformatics: Second Edition. Springer, Dordrecht, 3–35.

Lee SW, Keeney DR, Lim DH, Dispirito AA, Semrau JD. 2006. Mixed pollutant degradation by *Methylosinus trichosporium* OB3b expressing either soluble or particulate methane monooxygenase: Can the tortoise beat the hare. *Applied and Environmental Microbiology* **72**, 7503–7509.

Lee H, Sparkes I, Gattolin S, Dzimitrowicz N, Roberts LM, Hawes C, Frigerio L. 2013. An Arabidopsis reticulon and the atlastin homologue RHD3-like2 act together in shaping the tubular endoplasmic reticulum. *New Phytologist* **197**, 481–489.

Lieberman RL, Rosenzweig AC. 2004. Biological methane oxidation: regulation, biochemistry, and active site structure of particulate methane monooxygenase. *Critical Reviews in Biochemistry and Molecular Biology* **39**, 147–164.

Lieberman RL, Rosenzweig AC. 2005. Crystal structure of a membrane-bound metalloenzyme that catalyses the biological oxidation of methane. *Nature* **43**, 177–182.

Lieberman RL, Shrestha DB, Doan PE, Hoffman BM, Stemmler TL, Rosenzweig AC. 2003. Purified particulate methane monooxygenase from *Methylococcus capsulatus* (Bath) is a dimer with both mononuclear copper and a copper-containing cluster. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 3820–3825.

- Lin Y, Hung C-Y, Bhattacharya C, Nichols S, Rahimuddin H, Kittur FS, Leung T, Xie J.** 2018. An effective way of producing fully assembled antibody in transgenic tobacco plants by linking heavy and light chains via a self-cleaving 2A peptide. *Frontiers in Plant Science* **9**, 1379.
- Lin C, White RR, Sparkes I, Ashwin P.** 2017. Modeling endoplasmic reticulum network maintenance in a plant cell. *Biophysical Journal* **113**, 214–222.
- Lin C, Zhang Y, Sparkes I, Ashwin P.** 2014. Structure and dynamics of ER: minimal networks and biophysical constraints. *Biophysical Journal* **107**, 763.
- Liu Z, Chen O, Wall JBJ, Zheng M, Zhou Y, Wang L, Ruth Vaseghi H, Qian L, Liu J.** 2017. Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Scientific Reports* **7**, 1–9.
- Liu M, Lu S.** 2016. Plastoquinone and ubiquinone in plants: biosynthesis, physiological function and metabolic engineering. *Frontiers in Plant Science* **7**, 1898.
- Liu H, Zhang H, Wang G, Shen Z.** 2008. Identification of rice varieties with high tolerance or sensitivity to copper. *Journal of Plant Nutrition* **31**, 121–136.
- Lombardi R, Circelli P, Villani ME, Buriani G, Nardi L, Coppola V, Bianco L, Benvenuto E, Donini M, Marusic C.** 2009. High-level HIV-1 Nef transient expression in *Nicotiana benthamiana* using the P19 gene silencing suppressor protein of Artichoke Mottled Crinckle Virus. *BMC Biotechnology* **9**, 1–11.
- Lontoh S, Semrau JD.** 1998. Methane and trichloroethylene degradation by *Methylosinus trichosporium* OB3b expressing particulate methane monooxygenase. *Applied and Environmental Microbiology* **64**, 1106–1114.
- Lössl A, Eibl C, Harloff HJ, Jung C, Koop HU.** 2003. Polyester synthesis in transplastomic tobacco (*Nicotiana tabacum* L.): Significant contents of polyhydroxybutyrate are associated with growth reduction. *Plant Cell Reports* **21**, 891–899.
- Lu YJ, Hung MC, Chang BTA, et al.** 2019. The PmoB subunit of particulate methane monooxygenase (pMMO) in *Methylococcus capsulatus* (Bath): The Cu I sponge and its function. *Journal of Inorganic Biochemistry* **196**.
- Luginbuehl LH, El-Sharnouby S, Wang N, Hibberd JM.** 2020. Fluorescent reporters for functional analysis in rice leaves. *Plant Direct* **4**, e00188.
- Luke GA, Roulston C, Tilsner J, Ryan MD.** 2015. Growing uses of 2A in plant biotechnology. *Biotechnology*, 165-193.

Luo L, Zhou WQ, Liu P, Li CX, Hou SW. 2012. The development of stomata and other epidermal cells on the rice leaves. *Biologia Plantarum* 2012 56:3 **56**, 521–527.

Mackay JP, Landsberg MJ, Whitten AE, Bond CS. 2017. Whaddaya know: a guide to uncertainty and subjectivity in structural biology. *Trends in Biochemical Sciences* **42**, 155–167.

Masson-Delmotte, V., P. Zhai, A. Pirani SL, Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy JBR, Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu and BZ. 2021. *IPCC, 2021: Climate Change 2021: The Physical Science Basis*.

McGiffen ME, Manthey JA. 1996. The role of methanol in promoting plant growth: A current evaluation. *HortScience* **31**, 1092–1096.

Le Mer J, Roger P. 2001. Production, oxidation, emission and consumption of methane by soils: A review. *European Journal of Soil Biology* **37**, 25–50.

Minamikawa K, Sakai N, Yagi K. 2006. Methane emission from paddy fields and its mitigation options on a field ccale. *Microbes Environ* **21**, 135–147.

Minskaia E, Ryan MD. 2013. Protein coexpression using FMDV 2A: effect of “linker” residues. *BioMed Research International* **2013**, 1–12.

Mitić N, Schwartz JK, Brazeau BJ, Lipscomb JD, Solomon EI. 2008. CD and MCD studies of the effects of component B variant binding on the biferrous active site of methane monooxygenase. *Biochemistry* **47**, 8386–8397.

Mitra K, Ubarretxena-Belandia I, Taguchi T, Warren G, Engelman DM. 2004. Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 4083.

Mitsuhashi N, Shimada T, Mano S, Nishimura M, Hara-Nishimura I. 2000. Characterization of organelles in the vacuolar-sorting pathway by visualization with GFP in tobacco BY-2 Cells. *Plant and Cell Physiology* **41**, 993–1001.

Miyaji A, Nitta M, Baba T. 2019. Influence of copper ions removal from membrane-bound form of particulate methane monooxygenase from *Methylosinus trichosporium* OB3b on its activity for methane hydroxylation. *Journal of Biotechnology* **306**, 100001.

Moeller L, Gan Q, Wang K. 2009. A bacterial signal peptide is functional in plants and directs proteins to the secretory pathway. *Journal of Experimental Botany* **60**, 3337–52.

Morris RH. 2020. Methane activation by a single copper center in particulate methane monooxygenase: A computational study. *Inorganica Chimica Acta* **503**, 119441.

Mourrain P, Béclin C, Elmayan T, et al. 2000. Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. *Cell* **101**, 533–542.

Murrell JC, Gilbert B, McDonald IR. 2000. Molecular biology and regulation of methane monooxygenase. *Archives of Microbiology* **173**, 325–332.

Napier RM, Fowke LC, Hawes C, Lewis M, Pelham HR. 1992. Immunological evidence that plants use both HDEL and KDEL for targeting proteins to the endoplasmic reticulum. *Journal of Cell Science* **102**.

Naqvi S, Farré G, Sanahuja G, Capell T, Zhu C, Christou P. 2010. When more is better: multigene engineering in plants. *Trends in Plant Science* **15**, 48–56.

Ngoi S, Chien A, Lee C. 2004. Exploiting internal ribosome entry sites in gene therapy vector design. *Current Gene Therapy* **4**, 15–31.

Nguyen HHT, Shiemke AK, Jacobs SJ, Hales BJ, Lidstrom ME, Chan SI. 1994. The nature of the copper ions in the membranes containing the particulate methane monooxygenase from *Methylococcus capsulatus* (Bath). *Journal of Biological Chemistry* **269**, 14995–15005.

Nielsen KA, Tattersall DB, Jones PR, Lindberg Møller B. 2007. Metabolon formation in dhurrin biosynthesis. *Phytochemistry* **69**, 88–98.

Nisbet EG, Dlugokencky EJ, Bousquet P. 2014. Methane on the rise-again. *Science, New Series* **343**, 493–495.

Nixon-Abell J, Obara CJ, Weigel A V., et al. 2016. Increased spatiotemporal resolution reveals highly dynamic dense tubular matrices in the peripheral ER. *Science* **354**.

Nouchi I, Mariko S, Aoki K. 1990. Mechanism of methane transport from the rhizosphere to the atmosphere through rice. *Physiology* **94**, 59–66.

Pachauri RK, Meyer L, Hallegatte France S, Bank W, Hegerl G, Brinkman S, van Kesteren L, Leprince-Ringuet N, van Boxmeer F. 2014. *IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Gian-Kasper Plattner.

Pagny S, Lerouge P, Ronique Gomord V. 1999. Signals and mechanisms for protein retention in the endoplasmic reticulum. *Journal of Experimental Botany*, **50**, 157–164.

Pain C, Kriechbaumer V. 2020. Defining the dance: quantification and classification of

endoplasmic reticulum dynamics. *Journal of Experimental Botany* **71**, 1757–1762.

Pain C, Kriechbaumer V, Kittelmann M, Hawes C, Fricker M. 2019. Quantitative analysis of plant ER architecture and dynamics. *Nature Communications* **10**, 1–15.

Paine JA, Shipton CA, Chaggar S, et al. 2005. Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nature Biotechnology* **23**, 482–487.

Peng W, Qu X, Shaik S, Wang B. 2021. Deciphering the oxygen activation mechanism at the CuC site of particulate methane monooxygenase. *Nature Catalysis* **4**, 266–273.

Pengelly JJJ, Förster B, von Caemmerer S, Badger MR, Price GD, Whitney SM. 2014. Transplastomic integration of a cyanobacterial bicarbonate transporter into tobacco chloroplasts. *Journal of Experimental Botany* **65**, 3071–80.

Peyret H, Lomonossoff GP. 2015. When plant virology met *Agrobacterium*: the rise of the deconstructed clones. *Plant Biotechnology Journal* **13**, 1121.

Philippot L, Hallin S. 2011. Towards food, feed and energy crops mitigating climate change. *Trends in Plant Science* **16**, 476–480.

Poulter B, Bousquet P, Canadell JG, et al. 2017. Global wetland contribution to 2000–2012 atmospheric methane growth rate dynamics. *Environmental Research Letters* **12**, 094013.

Rajeev Kumar S, Anunanthini P, Ramalingam S. 2015. Epigenetic silencing in transgenic plants. *Frontiers in Plant Science* **0**, 693.

Ralston L, Yu O. 2006. Metabolons involving plant cytochrome P450s. *Phytochemistry Reviews* **5**, 459–472.

Ramírez I, Dorta F, Espinoza V, Jiménez E, Mercado A, Peña-Cortés H. 2006. Effects of foliar and root applications of methanol on the growth of *Arabidopsis*, tobacco, and tomato plants. *Journal of Plant Growth Regulation* **25**, 30–44.

Rapoport TA. 2007. Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature* **450**, 663–669.

Ratajczak E, Petcherski A, Ramos-Moreno J, Ruonala MO. 2014. FRET-assisted determination of CLN3 membrane topology. *PLOS ONE* **9**, e102593.

Reinhardt A, Kagawa H, Woltjen K. 2020. N-terminal amino acids determine KLF4 protein stability in 2A peptide-linked polycistronic reprogramming constructs. *Stem Cell Reports* **14**, 520–527.

Renna L, Hanton SL, Stefano G, Bortolotti L, Misra V, Brandizzi F. 2005. Identification and characterization of AtCASP, a plant transmembrane Golgi matrix protein. *Plant Molecular Biology* 2005 58:1 **58**, 109–122.

Ro SY, Schachner LF, Koo CW, et al. 2019. Native top-down mass spectrometry provides insights into the copper centers of membrane-bound methane monooxygenase. *Nature Communications* **10**.

Rogelj J, Den Elzen M, Höhne N, Fransen T, Fekete H, Winkler H, Schaeffer R, Sha F, Riahi K, Meinshausen M. 2016. Paris Agreement climate proposals need a boost to keep warming well below 2 °C. *Nature* **534**, 631–639.

Ross MO, MacMillan F, Wang J, Nisthal A, Lawton TJ, Olafson BD, Mayo SL, Rosenzweig AC, Hoffman BM. 2019. Particulate methane monooxygenase contains only mononuclear copper centers. *Science* **364**, 566–570.

Ross MO, Rosenzweig AC. 2017. A tale of two methane monooxygenases. *Journal of Biological Inorganic Chemistry* **22**, 307–319.

Runions J, Brach T, Kühner S, Hawes C. 2006. Photoactivation of GFP reveals protein dynamics within the endoplasmic reticulum membrane. *Journal of Experimental Botany* **57**, 43–50.

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

Samalova M, Fricker M, Moore I. 2006. Ratiometric fluorescence-imaging assays of plant membrane traffic using polyproteins. *Traffic* **7**, 1701–1723.

Saunois M, Bousquet P, Poulter B, et al. 2016a. The global methane budget 2000–2012. *Earth System Science Data* **8**, 697–751.

Saunois M, Jackson RB, Bousquet P, Poulter B, Canadell JG. 2016b. The growing role of methane in anthropogenic climate change. *Environmental Research Letters* **11**, 120207.

Saxena B, Subramaniyan M, Malhotra K, Bhavesh NS, Potlakayala SD, Kumar S. 2014. Metabolic engineering of chloroplasts for artemisinin acid biosynthesis and impact on plant growth. *Journal of Biosciences* **39**, 33–41.

Scanlan J, Dumont MG, Murrell JC. 2009. Involvement of MmoR and MmoG in the transcriptional activation of soluble methane monooxygenase genes in *Methylosinus trichosporium* OB3b. *FEMS Microbiology Letters* **301**, 181–187.

Schoberer J, Botchway SW. 2014. Investigating protein-protein interactions in the plant endomembrane system using multiphoton-induced FRET-FLIM. *Methods in Molecular Biology*. Humana Press, New York, NY, 81–95.

Schroeder LK, Barentine AES, Merta H, Schweighofer S, Zhang Y, Baddeley D, Bewersdorf J, Bahmanyar S. 2019. Dynamic nanoscale morphology of the ER surveyed by STED microscopy. *Journal of Cell Biology* **218**, 83–96.

Schubert D, Lechtenberg B, Forsbach A, Gils M, Bahadur S, Schmidt R. 2004. Silencing in Arabidopsis T-DNA transformants: the predominant role of a gene-specific RNA sensing mechanism versus position effects. *The Plant Cell* **16**, 2561.

Schwarz DS, Blower MD. 2016. The endoplasmic reticulum: structure, function and response to cellular signaling. *Cellular and Molecular Life Sciences* **73**, 79–94.

Semrau JD. 2011. Bioremediation via methanotrophy: overview of recent findings and suggestions for future research. *Frontiers in Microbiology* **2**.

Semrau JD, DiSpirito AA, Gu W, Yoon S. 2018. Metals and methanotrophy. *Applied and Environmental Microbiology* **84**.

Semrau JD, DiSpirito AA, Yoon S. 2010. Methanotrophs and copper. *FEMS Microbiology Reviews* **34**, 496–531.

Serrano-Silva N, Sarria-Guzmán Y, Dendooven L, Luna-Guido M. 2014. Methanogenesis and methanotrophy in soil: A Review. *Pedosphere* **24**, 291–307.

Setyanto P, Rosenani AB, Boer R, Fauziah CI, Khanif MJ. 2004. The effect of rice Cultivars on methane emission from irrigated rice field. *Indonesian Journal of Agricultural Science* **5**, 20.

Shanmugaraj B, Bulaon CJI, Phoolcharoen W. 2020. Plant molecular farming: A viable platform for recombinant biopharmaceutical production. *Plants* **9**, 1–19.

Shiemke AK, Cook SA, Miley T, Singleton P. 1995. Detergent solubilization of membrane-bound methane monooxygenase requires plastoquinol analogs as electron donors. *Archives of Biochemistry and Biophysics* **321**, 421–428.

Silva-Alvim FAL, An J, Alvim JC, Foresti O, Grippa A, Pelgrom A, Adams TL, Hawes C, Denecke J. 2018. Predominant golgi residency of the plant K/HDEL receptor is essential for its function in mediating ER retention. *Plant Cell* **30**, 2174–2196.

Singh JS, Singh DP. 2017. Methanotrophs: an emerging bioremediation tool with unique broad spectrum methane monooxygenase (MMO) enzyme. *Agro-Environmental Sustainability*

2, 1–18.

Sirajuddin S, Barupala D, Helling S, Marcus K, Stemmler TL, Rosenzweig AC. 2014. Effects of zinc on particulate methane monooxygenase activity and structure. *The Journal of Biological Chemistry* **289**, 21782.

Sirajuddin S, Rosenzweig AC. 2015. Enzymatic oxidation of methane. *Biochemistry* **54**, 2283–94.

Smith SM, Rawat S, Telser J, Hoffman BM, Stemmler TL, Rosenzweig AC. 2011. Crystal structure and characterization of particulate methane monooxygenase from *Methylocystis* species Strain M. *Biochemistry* **50**, 10231–10240.

Sparkes I. 2018. Lessons from optical tweezers: quantifying organelle interactions, dynamics and modelling subcellular events. *Current Opinion in Plant Biology* **46**, 55–61.

Sparkes I, Runions J, Hawes C, Griffing L. 2009. Movement and remodeling of the endoplasmic reticulum in nondividing cells of tobacco leaves. *Plant Cell* **21**, 3937–3949.

Sparkes IA, Runions J, Kearns A, Hawes C. 2006. Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nature Protocols* **1**, 2019–2025.

Sparkes I, Tolley N, Aller I, Svozil J, Osterrieder A, Botchway S, Mueller C, Frigerio L, Hawes C. 2010. Five arabidopsis reticulon isoforms share endoplasmic reticulum location, topology, and membrane-shaping properties. *The Plant Cell* **22**, 1333–1343.

Spits M, Heesterbeek IT, Voortman LM, Akkermans JJ, Wijdeven RH, Cabukusta B, Neefjes J. 2021. Mobile late endosomes modulate peripheral endoplasmic reticulum network architecture. *EMBO reports* **22**, e50815.

Strand S, Bruce N, Rylott L, Zhang L. 2013. *Phytoremediation of Atmospheric Methane*.

Stubbs CD, Botchway SW, Slater SJ, Parker AW. 2005. The use of time-resolved fluorescence imaging in the study of protein kinase C localisation in cells. *BMC Cell Biology* **2005** 6:1 **6**, 1–12.

Su J, Hu C, Yan X, et al. 2015. Expression of barley SUSIBA2 transcription factor yields high-starch low-methane rice. *Nature* **523**, 602–606.

Sun J, Movahed N, Zheng H. 2020. LUNAPARK is an E3 ligase that mediates degradation of ROOT HAIR DEFECTIVE3 to maintain a tubular ER network in Arabidopsis. *The Plant Cell* **32**, 2964–2978.

Sun H, Zhou S, Fu Z, Chen G, Zou G, Song X. 2016. A two-year field measurement of methane and nitrous oxide fluxes from rice paddies under contrasting climate conditions. *Scientific Reports* **6**, 28255.

Takeguchi M, Miyakawa K, Okura I. 1998. Properties of the membranes containing the particulate methane monooxygenase from *Methylosinus trichosporium* OB3b. *BioMetals* **11**, 229–234.

Tamura K, Shimada T, Ono E, Tanaka Y, Nagatani A, Higashi SI, Watanabe M, Nishimura M, Hara-Nishimura I. 2003. Why green fluorescent fusion proteins have not been observed in the vacuoles of higher plants. *Plant Journal* **35**, 545–555.

The C4 rice project. <https://c4rice.com/>

Ueda H, Yokota E, Kutsuna N, Shimada T, Tamura K, Shimmen T, Hasezawa S, Dolja V V, Hara-Nishimura I, Larkins BA. 2010. Myosin-dependent endoplasmic reticulum motility and F-actin organization in plant cells. **107**.

Urwin P, Yi L, Martin H, Atkinson H, Gilmartin PM. 2000. Functional characterization of the EMCV IRES in plants. *Plant Journal* **24**, 583–589.

Vaucheret H, Béclin C, Fagard M. 2001. Post-transcriptional gene silencing in plants. *Journal of Cell Science* **114**, 3083–3091.

Vaucheret H, Fagard M. 2001. Transcriptional gene silencing in plants: targets, inducers and regulators. *Trends in Genetics* **17**, 29–35.

Verdaasdonk JS, Lawrimore J, Bloom K. 2014. Determining absolute protein numbers by quantitative fluorescence microscopy. *Methods in cell biology* **123**, 347.

Verma D, Daniell H. 2007. Chloroplast vector systems for biotechnology applications. *Plant Physiology* **145**, 1129–1143.

Verma D, Samson NP, Koya V, Daniell H. 2008. A protocol for expression of foreign genes in chloroplasts. *Nature Protocols* **3**, 739–758.

Vitale A, Pedrazzini E. 2005. Recombinant pharmaceuticals from plants: the plant endomembrane system as bioreactor. *Molecular interventions* **5**, 216–225.

Voeltz GK, Prinz WA, Shibata Y, Rist JM, Rapoport TA. 2006. A class of membrane proteins shaping the tubular endoplasmic reticulum. *Cell* **124**, 573–586.

Voinnet O, Rivas S, Mestre P, Baulcombe D. 2003. An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus.

Plant Journal **33**, 949–956.

Wallrabe H, Periasamy A. 2005. Imaging protein molecules using FRET and FLIM microscopy. *Current Opinion in Biotechnology* **16**, 19–27.

Wang L, Liu Z, Yin C, Asfour H, Chen O, Li Y, Bursac N, Liu J, Qian L. 2015*a*. Stoichiometry of Gata4, Mef2c, and Tbx5 influences the efficiency and quality of induced cardiac myocyte reprogramming. *Circulation Research* **116**, 237–244.

Wang Y, Wang F, Wang R, Zhao P, Xia Q. 2015*b*. 2A self-cleaving peptide-based multi-gene expression system in the silkworm *Bombyx mori*. *Scientific reports* **5**, 16273.

Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S. 2011. A modular cloning system for standardized assembly of multigene constructs. *PLOS ONE* **6**, e16765.

Wells JN, Bergendahl LT, Marsh JA. 2016. Operon gene order is optimized for ordered protein complex assembly. *Cell Reports* **14**, 679–685.

Xing S, Wallmeroth N, Berendzen KW, Grefen C. 2016. Techniques for the analysis of protein-protein interactions *in vivo*. *Plant Physiology* **171**, 727–758.

Yao J, Weng Y, Dickey A, Wang KY. 2015. Plants as factories for human pharmaceuticals: applications and challenges. *International Journal of Molecular Sciences* **16**, 28549.

Ye X, Al-Babili S, Klöti A, Zhang J, Lucca P, Beyer P, Potrykus I. 2000. Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science (New York, N.Y.)* **287**, 303–5.

Yruela I. 2009. Copper in plants: acquisition, transport and interactions. *Functional Plant Biology* **36**, 409–430.

Yurimoto H, Shiraishi K, Sakai Y. 2021. Physiology of methylotrophs living in the phyllosphere. *Microorganisms* **9**.

Zahn JA, Dispirito AA. 1996. Membrane-associated methane monooxygenase from *Methylococcus capsulatus* (Bath). *Journal of Bacteriology* **178**, 1018–1029.

Zang J, Klemm S, Pain C, Duckney P, Bao Z, Stamm G, Kriechbaumer V, Bürstenbinder K, Hussey PJ, Wang P. 2021. A novel plant actin-microtubule bridging complex regulates cytoskeletal and ER structure at ER-PM contact sites. *Current Biology* **31**, 1251-1260.e4.

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

Zhu Q, Yu S, Zeng D, *et al.* 2017. Development of “Purple Endosperm Rice” by engineering anthocyanin biosynthesis in the endosperm with a high-efficiency transgene stacking system. *Molecular Plant* **10**, 918–929.