Determining targeting specificity of nuclearly encoded organelle proteins

Dissertation

zur Erlangung des
Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät I – Biowissenschaften –

der Martin-Luther-Universität
Halle-Wittenberg,

vorgelegt

von Herrn Mayank Sharma

geb. am 08.01.1990 in Mathura, Indien

Reviewers (Gutachter)

1. Prof. Dr. Ralf Bernd Klösgen
Institute of Biology-Plant Physiology, Martin Luther University, Halle (Saale) Germany

2. Prof. Dr. Bettina Hause
Leibniz Institute of Plant Biochemistry, Department of Metabolic and Cell Biology Halle (Saale) Germany

3. Prof. Dr. Michael Schroda
Molecular Biotechnology and Systems Biology, Kaiserslautern, Germany

Date of defense (Datum der öffentlichen Verteidigung):

06.06.2019

Table of Contents

List of abbreviations ..... **VI**

Glossary ..... **IX**

Summary **1**

Zusammenfassung **3**

Chapter 1 General Introduction

1.1 Endosymbiotic organelles of plant cells **5**

1.2 Protein transport into endosymbiotic organelles **6**

1.2.1 Protein targeting signals 6

1.2.2 Mitochondria protein import machinery 6

1.2.3 Plastid protein import machinery 8

1.3 Dual protein targeting to mitochondria and plastids **9**

1.3.1 Two principles of dual targeting 9

1.3.2 Proposed classification of dually targeted proteins with ambiguous transit peptides 10

1.4 Approaches to define protein targeting specificity **11**

1.4.1 Cell biology and biochemical assays 11

1.4.2 *In silico* targeting prediction 11

1.4.3 *In organello* protein transport experiments 12

1.4.4 Fluorescence microscopy based approaches to study protein targeting *in vivo* 13

1.5 Objectives of this work **16**

Chapter 2 Comparative analysis of fluorescence microscopy-based approaches to study targeting specificity of nuclearly encoded organelle proteins

2.1 Background **17**

2.2 Results **18**

2.2.1 Experimental set-up 18

2.2.2 Experimental methods 19

2.2.3 Subcellular localization of candidate proteins 19

(a) Mitochondrial Rieske protein and ferredoxin-NADP+-oxidoreductase (mtRi and FNR) 19

(b) Glycine cleavage system subunit-H 1 (GCS-H1) 21

(c) Co-chaperone/nucleotide exchange factor mitochondrial GrpE I (Mge1 or GrpE) 23

(d) Mitochondrial elongation factor Tu (EF-Tu) 25

(e) Peptide deformylase 1 B (PDF-1B or PDF) 26

2.3 Discussion **28**

2.3.1 Transient vs. Stable Transformation 28

2.3.2 Protoplasts vs. Intact Tissue 29

2.3.3 Presence or Absence of *Agrobacterium* in the Assays 30

2.3.4 Choice of the Experimental System for Protein Transport Studies 30

Chapter 3 The self-assembling split-fluorescent protein system to determine targeting specificity of nuclearly encoded organelle proteins

3.1 Background **32**

3.2 Results **34**

3.2.1 The *sa*split-GFP system to determine protein targeting specificity 34

3.2.2 Multimerization of the GFP11 tag leads to fluorescence signal enhancement in plastids but not in mitochondria 36

3.2.3 Construction of the PlaMiNGo toolkit 38

3.2.4 Multicolor imaging of dual protein targeting to two organelles 41

3.2.5 Analysis of protein targeting specificity with PlaMiNGo 43

3.2.6 Phenotype complementation confirms the plastid targeting properties of nuclearly encoded proteins 46

3.3 Discussion **49**

3.3.1 Selective imaging and fluorescence signal enhancement with sasplit-GFP technology 49

(a) Selective Imaging 49

(b) Fluorescence signal enhancement 49

3.3.2 High sensitivity of *sa*split-fluorescence protein system 50

3.3.3 Modularity of PlaMiNGO toolkit 51

3.3.4 Phenotype complementation as an alternative approach? 51

Chapter 4 Determining targeting specificity of chloroplast twin arginine translocase subunit A (TatA)

4.1 Background **53**

4.2 Results **55**

4.2.2 Subcellular localization of *Arabidopsis thaliana* TatA 55

4.2.3 Subcellular localization of *Pisum sativum* (pea) TatA 57

4.2.4 Determining mitochondrial targeting specificity of *Pisum sativum* (pea) TatA with *sa*split-GFP system 59

4.2.5 Generation of TatA *knock-out* mutant of *Arabidopsis thaliana* 59

4.2.6 Ultrastructure of subcellular compartments in TatA *knock-out* mutant 61

4.3 Discussion **63**

Chapter 5 General Discussion

5.1 Experimental approaches to define dual protein targeting 65

5.2 Evolutionary significance of dual protein targeting 66

5.3 Functional relevance of dual protein targeting 68

5.4 Perspectives 69

Chapter 6 Materials and Methods

6.1 Materials 70

6.1.1 Chemicals 70

6.1.2 Antibiotics 70

6.1.3 Enzymes 70

6.1.4 Molecular Biology Kits 71

6.1.5 Vectors (Molecular Biology) 71

6.1.6 Oligonucleotides (Oligos) 71

6.1.7 Miscellaneous 73

6.1.8 Biological Material 73

6.1.9 Culture Media 74

6.2 Methods 74

6.2.1 Molecular Cloning 74

(a) eYFP reporter constructions (Chapter 2) 74

(b) Generation of sasplit-GFP vectors (Chapter 3) 74

(d) Construction of the PlaMiNGo toolkit (Chapter 3) 75

(d) Cloning of candidate proteins into PlaMiNGo vectors (Chapter 3) 75

(e) Cloning for complementation of the immutans phenotype (Chapter 3) 75

(f) Construction of Crispr/Cas9 vector and generation of A. thaliana tatA mutants (Chapter 4) 76

(g) Standard Golden Gate *Reaction* 76

(h) Vectors newly generated in this study 77

6.2.1 Cultivation of Plants 80

(a) Arabidopsis thaliana (Col-0 background) 80

(b) Arabidopsis thaliana (Immutans mutants) 80

(c) Nicotiana benthamiana 81

(d) Pisum sativum 81

6.2.2 Transient transformation 81

(a) Particle Bombardment 81

(b) Protoplast Transformation 81

(c) Agrobacterium mediated transient transformation of Nicotiana benthamiana.....82
6.2.3 Generation of *Arabidopsis* thaliana stable transgenic lines 82
6.2.5 DNA Extraction.....82
6.2.4 Confocal laser scanning microscopy (CLSM).....82
6.2.5 Image processing and signal quantification.....83
6.2.4 Western blot (Immunoblot).....83
References 84
Notes 98
Appendix..... I
Appendix-I Gene sequences I
Appendix-II Step-by-step protocols V
(A) Tape-Sandwich method for protoplast isolation..... V
(B) Cloning of fragments from pRT100mod vectors to pLSU4GG VII
Acknowledgements- VIII
Curriculum Vitae..... IX
Erklärung zur Dissertation..... X

List of abbreviations

AA	Amino acids
Agroinfiltration	*Agrobacterium* infiltration
Amp	Ampicillin
ATP	Adenosin-5´-triphosphat
ATPase	Adenosin triphosphatase
BASTA	A brand name for Glufosinate or phosphinothricin
BSA	Bovine serum albumin
C-Terminal	Carboxy-terminal of a polypeptide
CaMV	Cauliflower mosaic virus
CaMV 35S Promoter	Promoter of 35S RNA coding region in CaMV genome
ccdB	Toxic component of a type II toxin-antitoxin (TA) system; inhibitor of DNA gyrase
CLSM	Confocal laser scanning microscopy
Cytc1	Cytochrom c1-Protein from Cytochrom bc1-complex of mitochondria
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol
E. coli	*Escherichia coli*
e.g.	exempli gratia (Latin); for example
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
Ef-Tu	Elongation factor thermo unstable (mitochondria)
eGFP	enhanced Green Fluorescent Protein
eYFP	enhanced Yellow Fluorescent Protein
FNR	Ferredoxin—NADP(+) reductase
FP	Fluorescent protein
GCS	H-Protein of Glycine cleavage system
Gen	Gentamycin
GFP	Green Fluorescent Protein
GFP1-10	N-terminal 10 β-sheets of GFP
GFP1-10 OPT	soluble optimized GFP1-10 (Cabantous et al., 2005)
GFP11	C-terminal 11th β-sheet of GFP
GG	Golden Gate
GrpE	Gro P like gene E I (mitochondria)
HSP	Heat Shock Protein

Hyg	Hygromycin
i.e.	*id est (Latin);* that is
IM	Inner membrane
im	immutans mutant of Arabidopsis thaliana
Kan	Kanamycin
LB	Luria-Bertani Medium
MES	2-(N-morpholino)ethanesulfonic acid
MLU	Martin Luther University Halle-Wittenberg
MPP	Mitochondria processing peptidase
MS Medium	Murashige and Skoog Medium
mtRi	mitochondrial Rieske Iron-Sulfur Protein
N-terminal	Amino-terminal of a polypeptide
Oxa1	Cytochrome oxidase assembly 1
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PDF	Peptide deformylase
PlaMiNGo	Plastid and/or Mitochondria targeted proteins N- terminally fused to GFP11 tags via Golden Gate cloning
pLaNGo	Plastid targeted proteins N- terminally fused to GFP11 tags via Golden Gate cloning
PMF	Proton motive force
pMiNGo	Mitochondria targeted proteins N- terminally fused to GFP11 tags via Golden Gate cloning
Preprotein	Precursor protein
PTOX	Plastid terminal oxidase
PVDF	Polyvinylidene fluoride
Rif	Rifampicin
RT	Room Temperature
Rubisco	Ribulose-1,5-bisphosphat-Carboxylase/Oxygenase
*sa*split-FP	Self-assembling split-fluorescent protein
*sa*split-GFP	Self-assembling split-Green fluorescent protein
*sa*split-YFP	Self-assembling split-yellow fluorescent protein
SDS	Sodium dodecyl sulfate
Sec	Secretory Pathway
Spec	Spectinomycin
SPP	Stroma processing peptidase

SRP	Signal recognition particle
SSU	Small subunit of Rubisco
T-DNA	Transfer DNA
Tat	Twin arginine translocase
Tic	Translocon at the inner envelope membrane of chloroplasts
Tim	Translocon of the inner mitochondrial membrane
tOCS	octopine synthase gene terminator from *Agrobacterium tumefaciens*
t35S	Terminator of 35S RNA coding region in CaMV genome
Tom	Translocon at the outer mitochondrial membrane
tRbcS	rbcS-E9 gene terminator from *Pisum sativum*
Tris	Tris(hydroxymethyl)aminomethane
tRNA	transfer RNA
TyrRS	Tyrosyl-tRNA Synthetases
v/v	Volume by volume
viz.	*videlicet* (Latin); namely
vs.	Versus
w/v	Weight by volume;
$\Delta\Psi$	Membrane Potential (delta Psi)

Glossary

Note- Definitions are in the context of this dissertation work

Autogenous theory: An evolutionary theory that explains the origin of eukaryotic cells. The theory argues that organelles of eukaryotic cell arose directly from a single prokaryote ancestor by infoldings of the prokaryote plasma membrane.

Dual targeting: The term is repeatedly used to describe the phenomenon of targeting of a single, nuclearly encoded protein to two distinct subcellular compartments.

Endosymbiotic theory: An evolutionary theory that explains the origin of eukaryotic cells. It proposes that two organelles of eukaryotic cell i.e. plastids and mitochondria are evolved from free-living prokaryotes via two consecutive and independent endosymbiotic events in the host cell.

Fluorescent protein tagging: A term widely used to address an experimental approach that involves fusion of a fluorescent protein fragment to the candidate protein so that its subcellular location can be visualized via fluorescence microscopy after expression in a cell.

Golden Gate cloning: A molecular cloning method that can simultaneously and directionally assemble DNA fragments using TypeIIs restriction enzymes (in most instances BsaI and BpiI) and T4 DNA ligase.

Precursor protein (Preprotein): A nascent polypeptide chain translated by ribosome and is in the process of maturation, is termed as preprotein. In the context of protein transport, the term refers to the organelle targeted protein carrying a transport signal that is processed after protein transport into the respective organelle and thus generating the mature proteins.

Self-assembling split-GFP system: A modified fluorescent protein tagging system where a superfolder variable of GFP is split into two parts, a large fragment consist of initial 10 β-sheets and a smaller fragment representing 11th β sheet of GFP. These fragments are non-fluorescing alone but can generate a functional fluorophore via spontaneous assembly and without the need of any additional interacting partners. The term should not be confused with widely used 'split-GFP' for fluorescent complementation experiments.

Transit peptide: Amino-terminal polypeptide signal that is responsible for the transport of a nuclearly encoded protein into plastids or mitochondria. The term is used interchangeably with 'presequence' for proteins targeted to mitochondria.

Translocase: The term 'translocase' is often used interchangeably with 'translocon'. In particular, it specifically refers to translocons that facilitate energy driven protein transport.

Translocon: A multi-protein complex associated with the translocation of substrate proteins across the biological membranes.

Summary

The uniqueness of plant cells is marked by the presence of two organelles of endosymbiotic origin, namely mitochondria and chloroplasts (plastids) which have specific but partially overlapping functions, e.g. in energy metabolism of the cell. Despite housing residual genomes, most of their proteins are encoded in the nucleus, synthesized by cytosolic ribosomes and need to be "re"-transported into the respective target organelle. While transport is in most instances strictly monospecific, a group of proteins carries "ambiguous" transit peptides mediating transport into both, mitochondria and plastids (dual targeting). A number of different approaches including *in silico*, *in organello* as well as both transient and stable *in vivo* assays are established to determine the targeting specificity of such transit peptides. Due to variability in the results obtained from different experimental approaches, such dual targeting is often disputed particularly (even more) in those instances where the function of these proteins is enigmatic in one or the other organelle.

Fluorescence protein (FP)-tagging and subsequent *in vivo* protein localization is one such widely utilized experimental approaches. An assessment of the literature-derived dataset revealed that approximately 28% of total proteins analyzed via FP-tagging approach alone show variable targeting specificities in different independent studies. Therefore, in this work four commonly used *in vivo* approaches to determine subcellular protein localization (particle bombardment, protoplast transformation, agrobacterium infiltration, and stable transformation) were systematically compared. Interestingly, a given candidate protein does not always show the same targeting specificity in different assays indicating that neither of the experimental approaches reflects the organelle protein transport process in an unbiased manner. Instead, they all have their pros and cons. As shown here, the divergent or even contradictory results can still all be valid and true because each method can only shed light on a part of the whole transport process. Therefore, the choice of method is important and depends very much on the question to be addressed.

Furthermore, a novel *in vivo* method based on a self-assembling split-fluorescent protein (*sa*split-FP) system was established to circumvent some of the inherent drawbacks of the 'standard' FP-based approaches. With this system, it was now possible to detect dual targeting for a number of proteins, which had earlier been characterized as being targeted to a single organelle only. In a larger context, more than just increasing the list of such dually targeted proteins, this finding also highlights the evolutionarily conserved nature of organelle translocation machinery in plant cells. This widespread dual targeting phenomenon might also lead to transfer of a complete

biosynthetic pathway across the organelle borders or might facilitate the development of novel pathways in an organelle. The dual targeting characteristics of chloroplastic TatA presents one such example and supports the idea of the existence of a Tat machinery in mitochondria of higher plants.

Zusammenfassung

Pflanzenzellen sind durch das Vorhandensein von zwei Organellen endosymbiotischen Ursprungs gekennzeichnet, Mitochondrien und Chloroplasten (Plastiden). Beide Organellen erfüllen unterschiedliche, teilweise aber auch überlappende Funktionen in der Zelle, z.B. im Energiestoffwechsel. Trotz des jeweils vorhandenen Restgenoms sind die meisten ihrer Proteine im Zellkern codiert und müssen deshalb nach ihrer Synthese durch cytosolische Ribosomen in das jeweilige Zielorganell "zurück"transportiert werden. In der Regel erfolgt dieser Transport monospezifisch, d.h. nur in eines der Organellen. Allerdings konnten in den letzten Jahren einige Proteine identifiziert werden, die "mehrdeutige" Transitpeptide für den Transport sowohl in die Mitochondrien als auch in die Plastiden tragen. Ein solches *"dual targeting"* ist allerdings häufig umstritten, zum einen wegen der oftmals unklaren Funktion des Proteins in einem der Organellen, zum anderen wegen divergenter Ergebnisse bei Verwendung unterschiedlicher experimenteller Ansätze.

Zur Bestimmung der Transportspezifität solcher Organellproteine stehen verschiedene experimentelle Ansätze zur Verfügung. Vor allem das Fluoreszenzprotein (FP)-*tagging* und die anschließende mikroskopische Analyse der Proteinlokalisierung ist ein häufig verwendeter Ansatz. Allerdings zeigt die Auswertung der Literaturdaten, dass knapp ein Viertel aller Proteine, die ausschließlich mittels solcher FP-*tagging*-Versuche analysiert wurden, divergente Transportspezifitäten in verschiedenen unabhängigen Studien aufweisen. Deshalb sollten in dieser Arbeit vier häufig genutzte *in vivo*-Ansätze zur Proteinlokalisierung (biolistische Transformation, Protoplastentransformation, Agrobakterieninfektion, stabile Transformation) systematisch miteinander verglichen werden. Dabei konnte tatsächlich bestätigt werden, dass ein Kandidatenprotein in unterschiedlichen Ansätzen nicht immer die gleiche Organellspezifität aufweist, sondern manchmal ein stark abweichendes oder sogar widersprüchliches Transportverhalten zeigen kann. Allerdings zeigte sich auch, dass selbst widersprüchliche Ergebnisse korrekt sein können, da jede Methode immer nur einen bestimmten Teil des gesamten Transportvorgangs darstellen kann. Entscheidend für die Wahl des im Einzelfall am besten geeigneten experimentellen Ansatzes ist daher die genaue Fragestellung.

Einige Probleme solcher "klassischen" FP-basierten Versuche werden durch den Einsatz vollständiger Reporterproteine hervorgerufen, die sich oftmals schnell falten und dann nur noch eingeschränkt über Membranen transportiert werden können. Um dieses Problem zu umgehen, wurde im Rahmen dieser Arbeit erstmals eine *in vivo*-Methode auf Pflanzenzellen angewendet,

die auf spontan assemblierenden Teilen eines Fluoreszenzproteins (*self assembling split*-FP) beruht. Mit diesem System war es möglich, *dual targeting*-Verhalten selbst für eine Reihe von Proteinen nachzuweisen, die zuvor als eindeutig monospezifisch charakterisiert worden waren. Dadurch lässt sich nun aber nicht nur die Liste der *dual targeting*-Proteine erweitern, sondern es unterstreicht die evolutionäre Konservierung der Proteintransportapparate in den endosymbiontischen Organellen, die sogar verantwortlich für die Verschiebung ganzer Biosynthesewege zwischen den Organellen oder die Entwicklung neuer Wege sein kann. Ein Beispiel hierfür könnte das beobachtete *dual targeting*-Verhalten von TatA sein, dass das Vorhandensein einer potentiellen Tat-Maschinerie in pflanzlichen Mitochondrien nahelegt.

Chapter 1-
General Introduction

In 1938, French biologist Edouard Chatton proposed a division for all life forms on the basis of their subcellular organization, he termed as prokaryotes and eukaryotes (Sapp, 2005). This division was based on an observation that the genetic material of blue-green algae (eukaryote) is nucleated i.e. surrounded by a membrane while the genetic material of bacteria (prokaryote) is free-floating in the cytoplasm. Few decades later, Stanier and Niel (1962) confirmed this prokaryote-eukaryote dichotomy and reported supporting observations about this division. They highlighted that the eukaryotic cells are structurally complex living entities and consist of multiple subcellular compartments.

1.1 Endosymbiotic organelles of plant cells

A plant cell is a typical example of such extensive partitioning of the hydrophilic cytoplasm into separate, membrane-bound compartments (or cell organelles). The origin of these organelles fits well with the "Autogenous theory" of eukaryotic cell evolution (Taylor, 1976) with the exception of two double-membrane-bound organelles namely chloroplasts (plastids) and mitochondria. Interestingly, these organelles of plant cell are originated from free-living prokaryotes, which presumably were ancestors of recent α-proteobacteria and cyanobacteria, respectively. They were each engulfed in single and independent, even though sequential endosymbiotic events, by a (proto)eukaryotic host cell (Margulis, 1981). Their conversion from endosymbionts to organelles was accompanied by a massive transfer of genetic information from the ingested prokaryotes to the nuclear host genome (Palmer 1997; Adams et al. 2000). However, in neither case were all genes transferred. Consequently, both organelles still contain residual genomes which each comprise a small number of protein encoding genes, explicitly about 10-40 in mitochondria and approximately 100 in plastids (Gray, 1993). Considering the requirement of perhaps 2000 – 3000 protein species in either of these organelles (Van Wijk and Baginsky 2011; Rao et al., 2017) it becomes evident that a vast proportion of such organellar proteins needs to be imported from the cytosol. To facilitate transport of protein into these organelles, both mitochondria and plastids possess specific but largely similar translocation machinery.

1.2 Protein transport into endosymbiotic organelles

The resemblances between the protein translocation systems of plastids and mitochondria highlight a common evolutionary lineage of both organelles (Schleiff and Becker 2011), (Figure 1.1). Protein transport into these organelles takes place by combined efforts of transport signals, cytosolic guiding factors and multiprotein translocases. Proteins destined to organelles are often generated as precursor polypeptides carrying either a cleavable N-terminal transport signal (matrix and/or stroma targeted proteins) or a non-cleavable internal signal sequence (e.g. outer membrane proteins). These precursor proteins (preproteins) are maintained in an unfolded state by cytosolic chaperones, e.g. HSP70, HSP90 and 14-3-3 protein, prior to transport. The cytosolic factors also guide precursor proteins, based on their affinity, to organelle specific translocases (Guéra et al., 1993; Voisine et al., 1999; Agarraberes and Dice, 2001; Flores-Pérez and Jarvis 2013). There is enough literature available that demonstrates the functional and structural aspects of these organelle specific transport machineries and transport signals, however they are still not completely understood (recently reviewed in Bölter 2018, Ghifari et al., 2018 and Nakai, 2018).

1.2.1 Protein targeting signals

Majority of the nuclearly encoded organelle proteins, targeted to either mitochondrial matrix or plastid stroma, carries an amino-terminal cleavable transport signals addressed as transit peptide (plastids and mitochondria) or presequence (mitochondria). In general, these transit peptides/presequences are variable in length and do not have any overall consensus sequence or conserved motif. The plastid specific transit peptides are comparatively longer than mitochondria specific presequences with an average length of 51-60 AA and 21-30 AA respectively (Zhang and Glaser, 2002). These sequences are overall positively charged and have propensity to form amphiphilic α-helices (von Heijne et al., 1989). The mitochondria specific presequences are enriched in serine and arginine residues whereas, the plastid specific transit peptides are enriched in serine and hydroxylated amino acid residues, which in some cases are phosphorylated by cytosolic kinases (Bruce 2000; Martin et al., 2006).

1.2.2 Mitochondria protein import machinery

In higher plant mitochondria the receptor proteins at outer envelope membrane namely Tom20 and Tom9 interact with the presequence and direct preproteins towards the translocation pore of the outer membrane, formed by Tom40 (Jänsch et al., 1998; Schleiff and Turnbull 1998; Murcha et al., 2014). Presequence binding receptors and translocation pore together constitute the translocon of the outer mitochondrial membrane (Tom) complex (Figure 1.1). Preproteins, destined to mitochondrial matrix, are later forwarded to translocon of the inner mitochondrial membrane (Tim) complex while those destined to inter-membrane space (IMS) take divergent

transport paths, e.g. via MIA:ERV pathway (Ghifari et al., 2018). On the other hand, the integral inner membrane (IM) proteins could either take a direct translocation path via inter-membrane space (IMS), sidestepping the interaction with Tim complex (e.g. carrier proteins) or they are first transported to mitochondrial matrix and later translocated to inner membrane via Oxa1 or putative Tat pathway (e.g. Cytc1 and Rieske iron sulfur protein; Stuart, 2002; Carrie et al., 2016). The Tim complex is composed of two β-barrel proteins namely Tim23 and Tim17, which together constitute a Tim23:17 complex in association with several other Tim proteins. The Tim23:17 complex represents a highly dynamic structural moiety and serves as the connecting link between Tom and Tim (Donzeau et al., 2000; Murcha et al., 2003).

In model plant, *Arabidopsis thaliana* three iso-genes encode Tim23 and Tim17 proteins each, which are differentially expressed and regulate the translocation of specific substrate pre-proteins (Ghifari et al., 2018). An ATP-driven pre-sequence translocase-associated motor (PAM) remains associated with Tim23:17 complex and assists pulling and folding of proteins destined to mitochondrial matrix. Two other PAM associated proteins viz. mitochondrial HSP70 (mtHsp70) and GrpE I (MgeI) function as chaperones that assist post-translocation protein folding. After import of preprotein, the presequence is immediately processed by a heterodimeric mitochondrial processing peptidase (MPP) (Murcha et al., 2004).

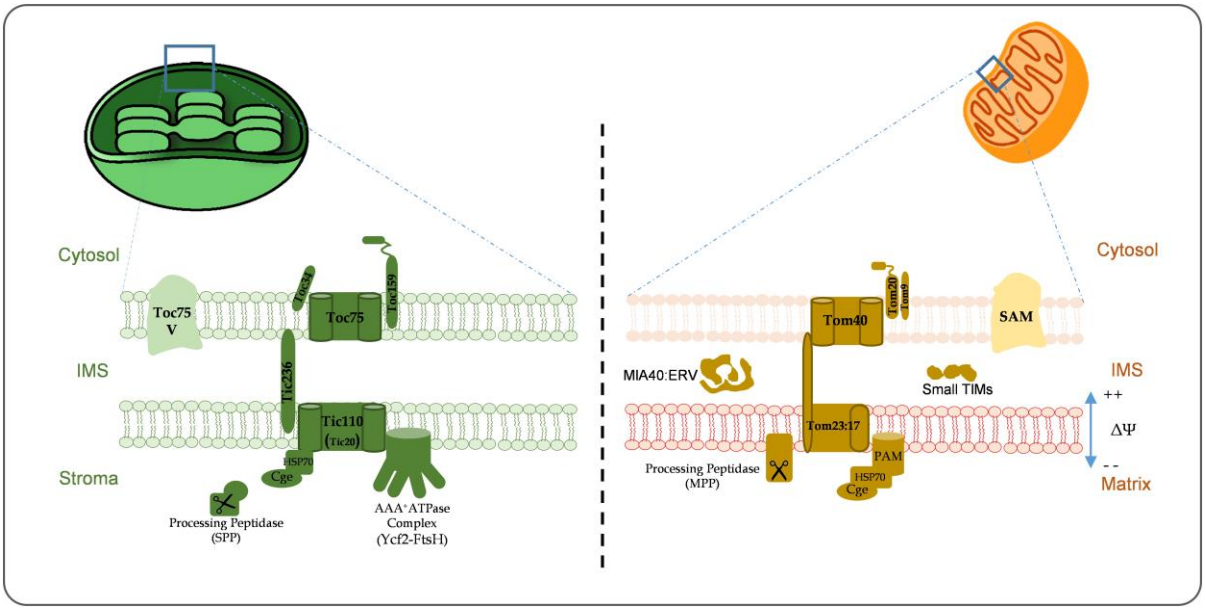


Figure 1.1 Simplified schematic representation of protein translocation machinery of plastids (*left*) and mitochondria (*right*)**.** The outer envelope translocon (Toc and Tom) of both organelles consists of preprotein receptors (Toc75, Toc34 and Tom20) and pore forming subunits (Toc75 and Toc40). Two additional translocons at outer membrane of plastids and mitochondria namely Toc75-V and SAM respectively, assist transport of outer envelope proteins. Plastid stroma and mitochondria matrix targeted proteins are further translocated via inner envelop translocons (Tic and Tim). The inner membrane translocons are consist of pore forming β-barrel protein subunits (Tic110 /Tic20 and Tim23/Tim17). In mitochondria, several translocons at IMS (MIA40:ERV and small TIMs) assist translocation of some of the proteins residing at IMS and envelop membranes. The ATP dependent protein pulling motors, namely AAA+ ATPase complex and PAM (presequence translocase-associated motor) complex in plastids and mitochondria respectively, provide a pulling force to preproteins and assist post-translocation protein folding. Membrane integrated processing peptidase of mitochondria (MPP) and stroma processing peptidase (SPP) facilitate cleavage of presequence/transit peptide. Toc/Tic- translocon at the outer/inner envelope membrane of chloroplast; Tom/Tim- translocon of the outer mitochondrial membrane; (*See text for more details*),

1.2.3 Plastid protein import machinery

In line with the mitochondrial protein import machinery, transport of the preproteins into plastid stroma is also facilitated by two multiprotein membrane complexes namely translocon at the outer envelope membrane of chloroplast (Toc) and translocon at the inner envelope membrane of chloroplast (Tic) (Figure 1.1). The GTP binding receptors, integrated in outer envelope, namely Toc159 and/or Toc34 recognizes the N-terminal transit peptides of preproteins (Fulgosi and Soll, 2002). Several homologs of these Toc receptors, e.g. Toc132, Toc120, Toc90 and Toc33, are differentially expressed in plant cell and influence the tissue/developmental specific import of specific preproteins (Richardson et al., 2017). The Toc receptors further assist the recruitment of preprotein to a cation selective channel, Toc75 that forms a β-barrel in the outer envelope membrane (Hinnah et al., 1997; 2002). Majority of plastid targeted proteins, except those targeted to the inter-membrane space e.g. Tic22 (Vojta et al., 2007), are then directed towards Tic complex and consecutively translocated to plastid stroma. After import into stroma, the transit peptide is processed by a stromal processing peptidase (SPP) (Robinson and Ellis, 1984).

The constitution and working model of translocon at the inner envelope membrane of chloroplast (Tic) remained controversial for many years and is still poorly understood (reviewed in Nakai, 2018). The classical models of Tic complex propose, undisputedly, Tic110/Tic40 and/or Tic20 as core components while Tic22, Tic56 and stromal chaperones as associated proteins. Additionally, three regulatory proteins namely Tic55, Tic62 and, Tic32 (constituting a putative redox regulon) were shown to control protein import by sensing the chloroplast redox status (Stengel et al., 2009). A recently characterized subunit, Tic236, is a novel addition to the existing Tic complex and proposed to serve as a connecting link between Toc and Tic (Chen et al., 2018). Likewise, a 2-megadalton heteromeric AAA^{+}-ATPase complex has been recently discovered and anticipated to be associated with Tic complex, working as a ATP driven preprotein pulling motor (Kikuchi et al., 2018).

In addition to envelope membranes, the chloroplasts (chlorophyll containing plastids) possess an internal system of interconnected membranes surrounding a lumen 'called' thylakoid. Protein transport into (or targeting to) thylakoids takes place via four major import pathways namely secretion (Sec) pathway, signal recognition particle (SRP) pathway, twin arginine translocation (Tat) pathway and spontaneous protein import pathway (reviewed in Gutensohn et al., 2006). Proteins destined to thylakoid compartment comprise a 'bipartite' transport signal that can be recognized by both Toc/Tic complex and by one of the thylakoid protein translocation system (Ko and Cashmore, 1989).

1.3 Dual protein targeting to mitochondria and plastids

The protein import machinery of both plastids and mitochondria has some profound differences but they are largely similar in terms of overall structure and preprotein passage. Likewise, these organelles possess many parallel biochemical pathways, for example, DNA replication, transcription, protein translation and electron transport. Because of these similarities, many proteins are inevitably required in both organelle and thus a large proportion of nuclearly encoded protein species is found in both mitochondria and plastids. Single genes in the nuclear genome encode majority of these proteins and consequently they are dually targeted to both endosymbiotic organelles. Huang and colleagues (1990) first coined the term 'dual targeting' in this context, as they were able to demonstrate dual targeting of a yeast mitochondrial signal peptide in isolated plant organelles. Enzymes involved in organelle protein synthesis, e.g. aminoacyl tRNA synthetases, and ribosomal proteins form a large proportion of these dually targeted proteins (listed in Carrie and Small, 2013; Carrie and Whelan, 2013).

1.3.1 Two principles of dual targeting

Closer inspection of genes encoding 'dually targeted' proteins reveals two different principles for how dual targeting can be achieved (Figure 1.2a). In the first one, the respective gene has the coding capacity for two distinct transport signals, which each mediate the specific transport into one of the organelles. Depending on mechanisms like alternative start of transcription or translation or alternative splicing either of the two transport signals can be N-terminally exposed after synthesis and thus determines the target organelle. Examples for such mechanisms have been described for protoporphyrinogen oxidase II (Watanabe et al., 2001), monodehydroascorbate reductase (Obara et al., 2002), THI1 protein (Chabregas et al., 2001), DNA polymerase POLγ2 (Christensen et al., 2005; Wamboldt et al., 2009), and glutamate receptor 3.5 (Teardo et al., 2015). The second principle describes dually targeted proteins that are derived from genes encoding single transport signals with 'ambiguous' targeting specificity (Figure 1.2b). The protein import machineries of both, mitochondria and chloroplasts recognize these 'so-called' dual targeting transit peptides. With respect to amino acid composition and structure, they share common features with "typical" transport signals for mitochondria as well as for chloroplasts (Ge et al., 2014). In line with that, they mediate transport of the passenger proteins by the same transport pathways that are used also by precursor proteins with single organelle specificity (Rödiger et al., 2010; Langner et al., 2014)

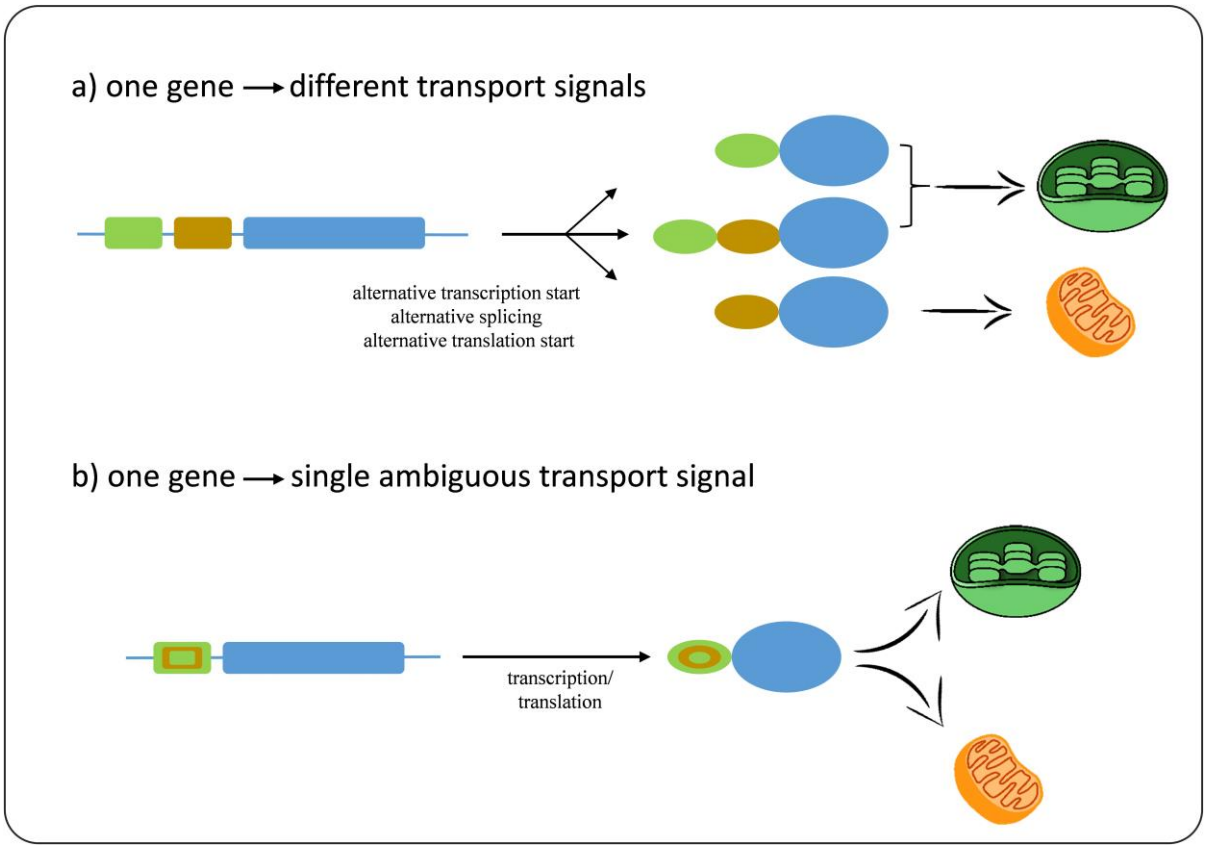


Figure 1.2 Two different principles of dual targeting. **a)** A single gene can have the coding capacity for two proteins carrying transit peptides with distinct organelle specificity which can be positioned at the N-terminus of the precursor protein as a result of alternative transcription start, alternative translation start and/or alternative exon splicing. **b)** The gene encodes a single precursor protein with an ambiguous transit peptide that mediates transport into both organelles. (*Source- Sharma et al., 2018*)

1.3.2 Proposed classification of dually targeted proteins with ambiguous transit peptides

Considering their functional relevance and the frequency of the corresponding genes in the nuclear genome, the precursor proteins comprising such dual targeting transit peptides can be further classified into three categories. Type I represents those proteins that have a proven function in both, mitochondria and chloroplasts and that are encoded by single nuclear genes. Typical examples for such proteins are tRNA synthetases, which are essential for protein synthesis in both organelles (Duchêne et al., 2005). Type II proteins are likewise required in both organelles but are encoded by small gene families. Only a subset of those proteins has dual targeting properties, while other members of the gene family might encode proteins with single organelle specificity. For example, among the four GrpE proteins encoded in the nuclear genome of *Arabidopsis thaliana*, two proteins each are addressed as mitochondrial (MgeI and MgeII) and plastidial components (CgeI and CgeII). Nevertheless, one member of the former group, namely MgeI, shows dual targeting into both organelles when analyzed by protein transport experiments (Hu et al., 2012; Baudisch et al., 2014). Dually targeted proteins of type III, on the other hand,

have a proven function in only one of the organelles, while its role in the other organelle remains enigmatic. One typical example is Cytochrome c_1, a component of the respiratory electron transport chain in mitochondria. This protein is transported also into chloroplasts (Rödiger et al., 2011), despite the fact that a function in this organelle appears rather unlikely.

1.4 Approaches to define protein targeting specificity

A common physiological role or an overlapping enzymatic function in mitochondria and chloroplasts often provides a first hint for potential dual targeting of the corresponding protein. However, to determine if a given protein is indeed imported into both endosymbiotic organelles, experimental verification is required for which a number of different assays are available.

1.4.1 Cell biology and biochemical assays

Probably the most direct approach is the isolation of intact organelles and screening for the occurrence of the candidate protein in the organellar fractions. In particular, screening of existing organelle proteomics data is a rapid and inexpensive method (Heazlewood, 2004; Sun et al., 2009). However, such data are available for only few plant species yet and their validity is strictly depending on the quality and purity of the organelles that were used for analysis. Furthermore, proteins of low abundance are difficult to identify with this approach. This holds true also if standard biochemical methods like Western analyses are employed. These often have the further constraint of lacking availability and/or specificity of antibodies, which likewise often prevents protein localization using immuno-gold labelling of plant tissue samples and subsequent visualization *via* electron microscopy. Considering furthermore the amount of labor and resources required for such experiments, neither of them appears suitable for screening of a large number of candidate proteins.

1.4.2 *In silico* targeting prediction

In contrast, computer-assisted prediction of protein targeting using online tools is an inexpensive and easy method to analyze a large number of proteins. Available information about N-terminal targeting signals, their amino acid composition, domain structure, sequence homology and/or gene ontology were used as training sets for the various prediction algorithms (Xiong et al., 2016). Several such tools are available nowadays. Among the most popular are TargetP, Predotar, WolfPsort, and MultiLoc (Emanuelsson et al., 2000; Small et al., 2004; Höglund et al., 2006; Horton et al., 2007). However, only some of them are actually adapted for the analysis of plant proteins. Furthermore, most of the algorithms were developed for proteins targeted to a single

organelle and are therefore not well suited to predict dually targeted proteins. For the analysis of such proteins, ATP (Ambiguous Targeting Predictor, Mitschke et al., 2009) and its modified version ATP2 (Fuss et al., 2013) were developed, which are based on experimentally validated positive and negative datasets of dually targeted proteins from several different plant species. As an alternative, SUBA, the subcellular database for Arabidopsis proteins has become popular, which is a manually curated dataset that is based upon the combination of experimental evidence (e.g. proteomics) and protein prediction tools (Hooper et al., 2017). However, all these prediction methods can only provide a kind of educated guess about the subcellular destination of a protein and require independent experimental verification.

1.4.3 *In organello* protein transport experiments

Different from the approaches described thus far, the following methods are based on experimental assays asking the question what happens if a candidate protein is presented to a potential target organelle. One such approach is an *in organello* protein transport experiment in which intact organelles that are freshly isolated from plant tissues are incubated with radiolabeled candidate precursor proteins obtained from *in vitro* translation in cell free systems like wheat germ extracts or reticulocyte lysates (Figure 1.3). The organelles can be isolated separately, e.g. if a mono-specific protein is under analysis, but, when analyzing potentially dually targeted proteins, simultaneous isolation of mitochondria and chloroplasts from a single pulping of plant tissue is probably more suitable (Rudhe et al., 2002; Rödiger et al., 2010). Organelles isolated from the same starting material abandon the possibility of differences in protein transport competence due to external variation, like for example growth conditions. Isolated intact organelles are physiologically active and contain the protein transport machinery facilitating the import of organelle-specific precursor proteins. However, besides the absence of cytosolic components, which obviously can play an important role in determining organelle specificity, these assays additionally inherit the problem of organellar cross-contamination. This holds true in particular for mitochondrial isolates being contaminated by proplastids, which are similar in mass and density (Keech et al., 2005). Since such cross-contamination cannot be completely prevented, *in organello* competition experiments are an appropriate complementation. In these experiments, organelle transport of a radiolabeled candidate protein is examined in the presence of excess amounts of an unlabeled precursor protein with known, and usually mono-specific, target organelle. Any significant decrease in organelle import of the candidate protein with increasing amounts of competitor confirms that the two proteins recognize the same organelle as target and utilize components of the same protein transport machinery (Rödiger et al., 2011; Langner et al., 2014).

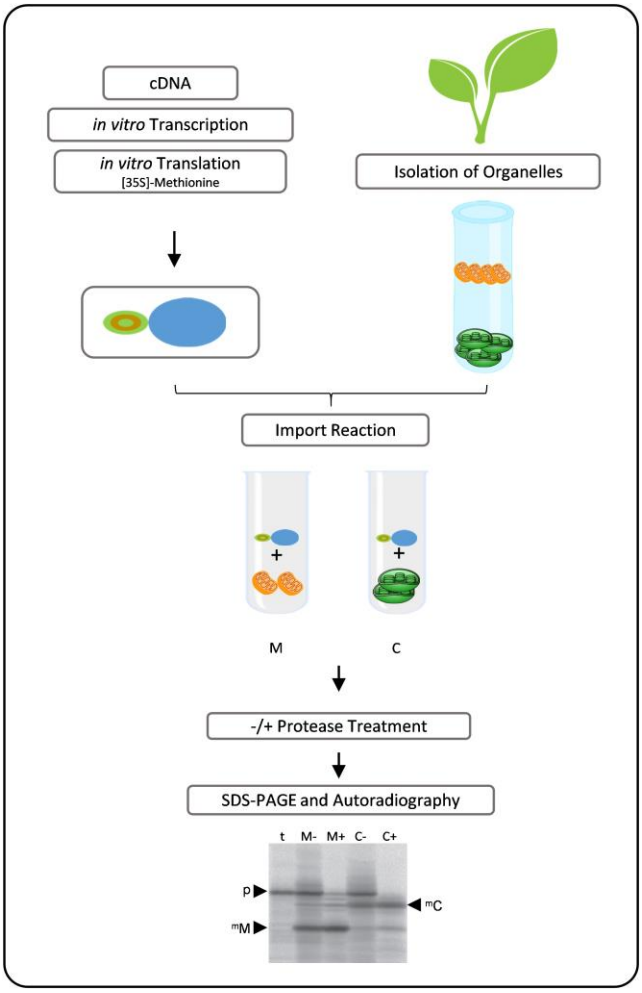


Figure 1.3 Workflow of an *in organello* import experiment with a dually targeted protein. Freshly isolated mitochondria (M) and chloroplasts (C) are incubated with the radiolabeled precursor protein (lane t) obtained from in vitro transcription and subsequent in vitro translation in the presence of [35S]-methionine. After incubation for 20 min in the light at 25 °C (Import Reaction), organelles are recovered and subsequently treated with either protease (lanes +) or mock-treated (lanes −). Stoichiometric amounts from each fraction are analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and detected by autoradiography. The bands representing the precursor protein (p) and the mature proteins accumulating in mitochondria (mM) and chloroplasts (mC) are indicated. (*Source- Sharma et al., 2018*)

1.4.4 Fluorescence microscopy based approaches to study protein targeting *in vivo*

The alternative approach to experimentally determine the organelle specificity of protein transport rests on live cell imaging employing fluorescent protein markers. In such assays, reporter constructs, which encode chimeric polypeptides comprising a fluorescent reporter protein that is C-terminally fused to a candidate transit peptide or full-length precursor, are transiently or stably introduced into living cells (Figure 1.4). After expression of such genes, the subcellular localization of the reporter can be traced by fluorescence microscopy. Amongst a wide variety of fluorescent reporters, the most popular for determining subcellular localization of a protein are GFP and YFP (green and yellow fluorescent proteins) (Sheen et al., 1995; Reichel et al., 1996). Their stability in folded state, long half-life time and high fluorescence intensity make them suitable fluorescent markers, in particular as slightly modified versions, namely enhanced GFP (eGFP) and YFP (eYFP) (Tsien, 1998). A variety of methods is available to transform plant cells with such reporter constructs. The quickest methods are those involving transient transformation

of plant tissues, namely (a) protoplast transformation of mesophyll cells, (b) biolistic transformation *via* particle bombardment, and (c) *Agrobacterium* infiltration of *Nicotiana benthamiana* leaves.

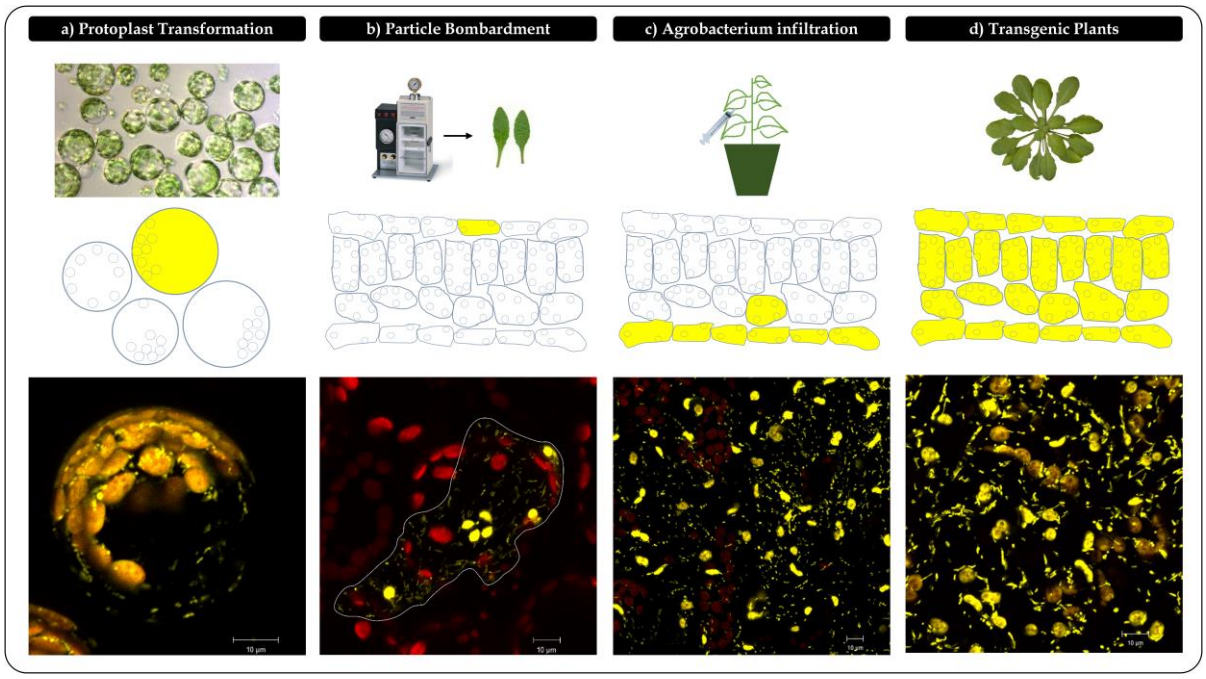


Figure 1.4 Approaches to study organelle targeting of proteins in vivo. **(a)** Transiently transformed mesophyll protoplasts surrounded by some untransformed protoplasts, **(b)** transiently transformed epidermal cell after particle bombardment, **(c)** *Agrobacterium* infiltration of *Nicotiana* epidermal cells, **(d)** leaf epidermis cells from a transgenic Arabidopsis line. The images were obtained from confocal laser scanning microscopy of cells expressing in all assays an identical reporter construct comprising an ambiguous transit peptide and eYFP. The yellow color represents signals obtained from eYFP fluorescence that is found in mitochondria (small dots) as well as in chloroplasts, which additionally show red chlorophyll autofluorescence. In the schematic drawings, the transformed cells/tissues are depicted in yellow color. (*Adapted from Sharma et al., 2018*)

The method of protoplast transformation using either PEG-mediated transformation or electroporation has successfully been developed already a while ago (Abel and Theologis, 1994). While being applicable for a wide variety of plant species including both monocots and dicots, the major disadvantage of this method is that, the cells are exposed to extreme stress during protoplast preparation (e.g., Papadakis and Roubelakis-Angelakis, 2002). It cannot be ruled out that this stress might affect the intracellular sorting of a protein. Particle bombardment, on the other hand, which can likewise be utilized to transform a variety of plant species, is generally used to transform epidermal cells of leaf tissue (Klein et al., 1992). Although the plastids found in such epidermal cells are usually small in size and do not resemble fully developed mesophyll chloroplasts, differences with regard to their properties in protein import have not yet been described (Dupree et al., 1991; Barton et al., 2016). The third method to transiently transform plant cells is based upon *Agrobacterium*-mediated transformation. This method is restricted to the

model plant genus *Nicotiana*, which consequently often confines the users to a heterologous system for their analysis. It is principally possible to infiltrate leaves of other plant species as well but it is apparently not as easy and efficient as with *Nicotiana* (Gelvin, 2003; Wroblewski et al., 2005). Furthermore, also in this case the plant cells experience considerable stress during transformation.

All three methods of transient transformation to analyze the organelle specificity of protein targeting are rather quick and reliable. However, they come with some obvious consequences that should be taken into the consideration. The number of genes expressed in the target cells and, thus, the amount of protein accumulating after expression cannot effectively be controlled and does usually not reflect the natural situation. Very high expression rates might even lead to the formation of aggresomes in the cytosol (García-Mata et al., 1999), which often resemble small organelles and are therefore difficult to distinguish from mitochondria after import of a fluorescent reporter protein (Toyooka et al., 2006). Therefore, the generation of stable transgenic lines is an alternative approach, which allows to focus the analysis to those plant lines with only a low copy number of transgenes in their nuclear genome. However, the method of stable transformation and regeneration of plants is still not established for a large number of plant species and is furthermore demanding with regard to time and labor.

1.5 Objectives of this work

The **(1)** first aim of this work was to systematically analyze four different *in vivo* experimental approaches (particle bombardment, protoplast transformation, *Agrobacterium* infiltration, and transgenic plants) that are commonly used to determine targeting specificity of nuclearly encoded plant proteins.

(2) The second aim of this dissertation was to establish alternative approaches to determine protein targeting specificity. This part involved critical assessment and optimization of self-assembling split-GFP (*sa*split-GFP) technology for its efficient utility in plant system.

(3) The third objective of this doctoral work was to determine targeting specificity of chloroplastic TatA from two plant model systems namely *Arabidopsis thaliana* and *Pisum sativum*.

Parts of this work have been published in the peer-reviewed journals with following titles-

(1) Sharma, M., Bennewitz, B., & Klösgen, R. B. (2018). Rather rule than exception? How to evaluate the relevance of dual protein targeting to mitochondria and chloroplasts. Photosynthesis Research. 138:335-343.

(2) Sharma, M., Bennewitz, B., & Klösgen, R. B. (2018). Dual or Not Dual? - Comparative Analysis of Fluorescence Microscopy-Based Approaches to Study Organelle Targeting Specificity of Nuclear-Encoded Plant Proteins. Frontiers in Plant Science. 9:1350.

(3) Sharma, M., Kretschmer, C., Lampe, C., Stuttmann, J., & Klosgen, R. B. (2019). Determining targeting specificity of nuclear-encoded organelle proteins with the self-assembling split fluorescent protein toolkit. Journal of Cell Science. 132:jcs230839.

Chapter 2- Comparative analysis of fluorescence microscopy-based approaches to study targeting specificity of nuclearly encoded organelle proteins

2.1 Background

Dual targeting of a protein to mitochondria and plastids is often claimed based on a single experimental approach (e.g. Huang et al., 1990; Creissen et al., 1995; Silva-Filho, 1999; Masuda et al., 2003; Carrie et al., 2008). However, reevaluation of these findings with alternative assays sometimes leads to contradicting results (e.g., Chow et al., 1997; Lister et al., 2001). This holds particularly true if principally different experimental approaches are compared with each other, like transport experiments with isolated organelles, microscopy of transiently or stably transformed plant tissue, biochemical assays or proteomics data (Tanz et al., 2013; Baudisch et al., 2014). While in *in vitro* assays, such as *in organello* protein transport experiments, the import of authentic precursor proteins into purified intact organelles is studied, *in vivo* approaches usually rest on the transient or stable expression of chimeric reporter constructs in living cells. Both approaches have their pros and cons and each can address only specific, and often different, aspects of the transport process. For example, *in organello* experiments examine if a given precursor protein is, in principle, a suitable substrate for the organellar import machinery. However, this approach does not clarify if such transport will actually take place also in the presence of potentially regulatory cytosolic factors present in intact cells. On the other hand, *in vivo* approaches often analyze the subcellular localization of chimeric proteins comprised of the candidate protein fused to a fluorescent reporter. The potential influence of the heterologous reporter on the transport process, for example, as a result of its folding properties or of its position within the chimera (e.g., either downstream of the transport signal or at the very C-terminus of the candidate protein) has been described (e.g., Marques et al., 2004; Baudisch et al., 2014) but was not systematically analyzed.

Surprisingly, even in those instances where largely similar approaches are applied, like in the various *in vivo* assays employing fluorescent reporter proteins, deviating results are sometimes described. For example, AOX2 (alternative oxidase-2) of *Arabidopsis thaliana* was described by Saisho et al. (2001) to be targeted solely to mitochondria and by Fu et al. (2012) to be targeted to chloroplasts. Another protein, OhmT (3-Methyl-2-oxobutanoate hydroxy-methyl-transferase) of *Arabidopsis thaliana*, showed with comparable *in vivo* approaches either monospecific transport into mitochondria (Ottenhof et al., 2004) or dual targeting to both endosymbiotic organelles (Baudisch et al., 2014). These contradicting results are often reported by separate research groups that might be using either different experimental protocols for a principally similar transformation approach or variable greenhouse conditions for growth of plants or different sets of chemicals, nevertheless, in all these instances the exact reason remains enigmatic. Therefore, the targeting behavior of four previously characterized dually targeted proteins was analyzed via four common *in vivo* experimental approaches and the results were compared to determine if there is variability in localization behavior of a protein due to differences in experimental assays.

2.2 Results

2.2.1 Experimental set-up

For the comparative analysis of the most common *in vivo* approaches used to study intracellular protein targeting in plants, we have focused on four candidate proteins from *Arabidopsis thaliana*, namely GCS (At2g35370), GrpE (At5g55200), EF-Tu (At4g02930), and PDF (At5g14660) (detailed description of proteins is available in the following sections). The proteins were originally selected from a systematic *in silico* prediction approach and all four candidates demonstrated dual targeting properties in *in organello* assays as well as in transiently transformed pea epidermal leaf cells (Baudisch et al., 2014). For comparison, two proteins with strictly monospecific organelle targeting characteristics (Rödiger et al., 2011) were analyzed in parallel, namely mtRi, the mitochondrial Rieske Fe/S protein from potato (GB:X79332.1) (Emmermann et al., 1994) and FNR (ferredoxin-NADP+-oxidoreductase, GB:M86349.1) from spinach (Zhang et al., 2001; Marques et al., 2004).

The reporter constructs as described in Baudisch et al. (2014), i.e., the N-terminal 100 amino acid residues of the respective precursor protein comprising the entire organelle targeting signal fused in-frame to eYFP (enhanced yellow fluorescent protein), were used in this analysis. The

exception was FNR, where the defined transit peptide of 55 amino acids was instead fused to eGFP (enhanced green fluorescent protein). This was not possible for the candidate proteins because for them neither the exact processing sites nor the actual targeting signals have yet been characterized in detail. It is important to note that in all our assays identical constructs were used. The constructs were cloned such that their expression in the plant cell is regulated by the CaMV 35S promoter and terminator.

2.2.2 Experimental methods

Four different experimental approaches were utilized to determine targeting specificity of candidate proteins. In the first set of *in vivo* experiments, the upper epidermis of *Arabidopsis thaliana* leaves were transiently transformed by (a) particle bombardment. This assay leads to the transformation of only a few, isolated cells. Transformed cells are distinguishable from the surrounding, non-transformed tissue by the accumulation of the fluorescent reporter protein, which is detected using epifluorescence microscopy. The second method was isolation and PEG mediated (b) transformation of *Arabidopsis thaliana* protoplasts. It requires more efforts than particle bombardment but has the advantage that a large number of cells can be transformed simultaneously. The third transient transformation system studied here is an inexpensive and technically rather simple assay involving (c) infiltration of *Agrobacterium* harboring the construct of interest in the lower epidermis of *Nicotiana benthamiana* leaves. Finally, (d) the stable transgenic lines of *Arabidopsis thaliana* were generated using the floral-dip method for comparison. This method is far more time-consuming than any of the transient transformation assays described above.

2.2.3 Subcellular localization of candidate proteins

(a) Mitochondrial Rieske protein and ferredoxin-NADP+-oxidoreductase (mtRi and FNR)

Two monospecific proteins, namely mtRi and FNR targeted to mitochondria and plastid respectively, were chosen as experimental control in this study. Rieske iron-sulphur protein is a part of cytochrome bc_1-complex (complex III) of the mitochondrial respiratory chain (Emmermann et al., 1994) while ferredoxin-NADP+-oxidoreductase catalyzes transfer of electron from ferredoxin to NADP+ during cyclic electron transport within chloroplast (Mulo, 2011). These proteins had previously been characterized as strictly monospecific proteins via *in organelle* import experiments and *in vivo* particle bombardment assays (Rödiger et al., 2011, Baudisch et al., 2014). When analyzed via four different experimental approaches, the expected

monospecific targeting to plastids and mitochondria was observed in majority of transformed cells expressing FNR_{1-55}/eGFP and $mtRi_{1-100}$/eYFP respectively (Figure 2.1, 2.2). The exception was protoplast transformation assays, where massive overexpression of reporter constructs resulted in artificial cytosolic localization of FNR_{1-55}/eGFP and aggregation of the chimeric $mtRi_{1-100}$/eYFP in few transformed protoplasts (Figure 2.1c, 2.2c).

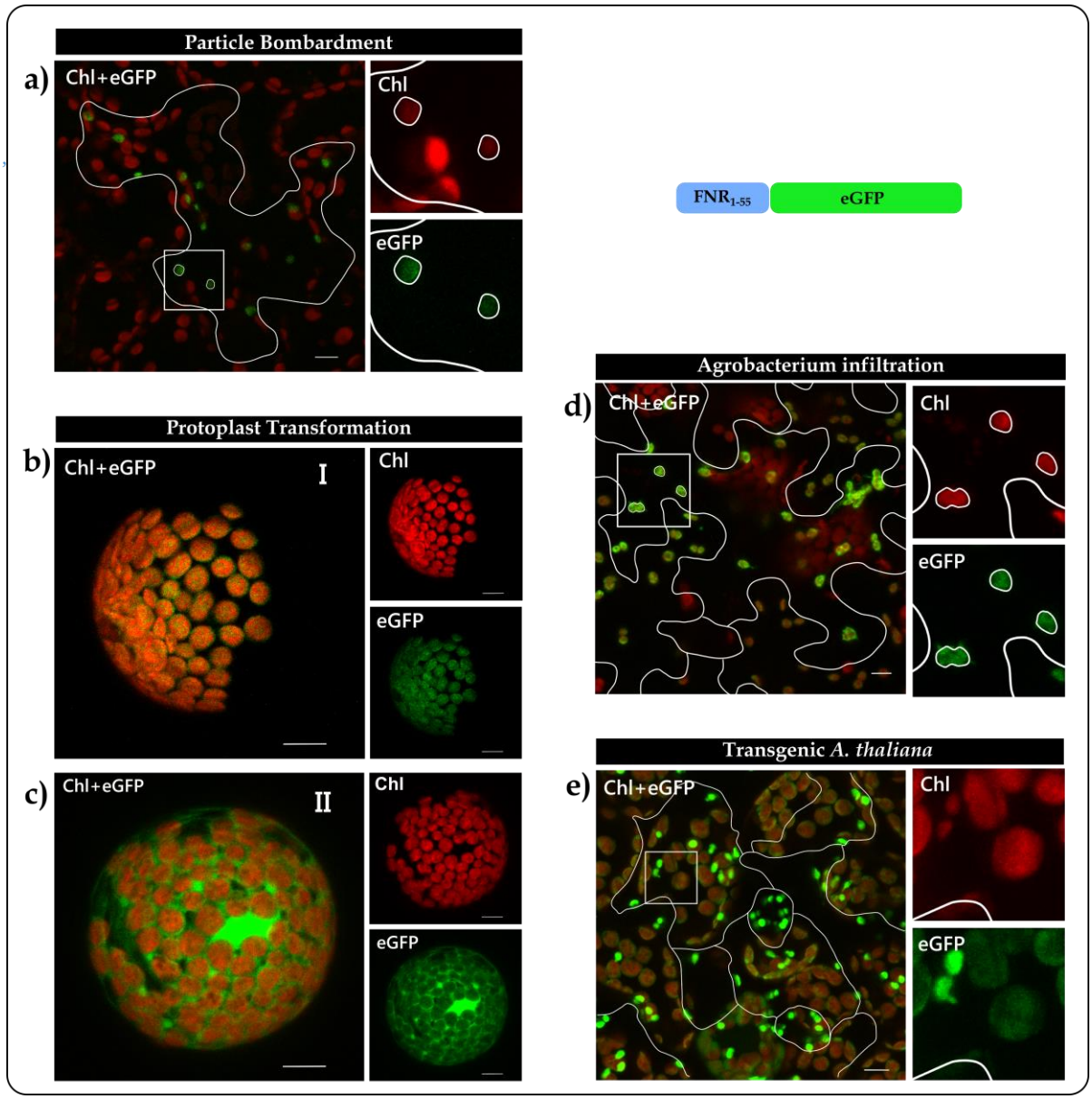


Figure 2.1 Subcellular localization of the candidate protein FNR_{1-55}/eGFP. The subcellular protein localization of FNR_{1-55}/eGFP was determined by confocal laser scanning microscopy after particle bombardment of leaf epidermal cells of *Arabidopsis thaliana* (panels **a**), protoplast transformation of *Arabidopsis thaliana* (panels **b** and **c**), *Agrobacterium* infiltration of *Nicotiana benthamiana* (panels **d**), and in transgenic *Arabidopsis thaliana* lines (panels **e**). Plastid localization of FNR_{1-55}/eGFP was observed in all instances. While, at high expression levels, FNR/eGFP is often found predominantly in the cytosol and nucleus rather than in plastids of transformed protoplast (panel **c**). Representative cells are displayed as overlay images of the chlorophyll channel (displayed in *red*) and the eGFP channel (displayed in *green*). In panels **a** and **d**, the strong chlorophyll signals in the background are derived from the larger chloroplasts of untransformed mesophyll cells underneath the epidermal cell layers. The squares highlight areas of the transformed cells that are shown in higher magnification separately for the chlorophyll channel and the eGFP channel, as indicated. All images are maximum intensity projections of several single images representing the complete cell in z-axis. The scale bars correspond to 10 µm.

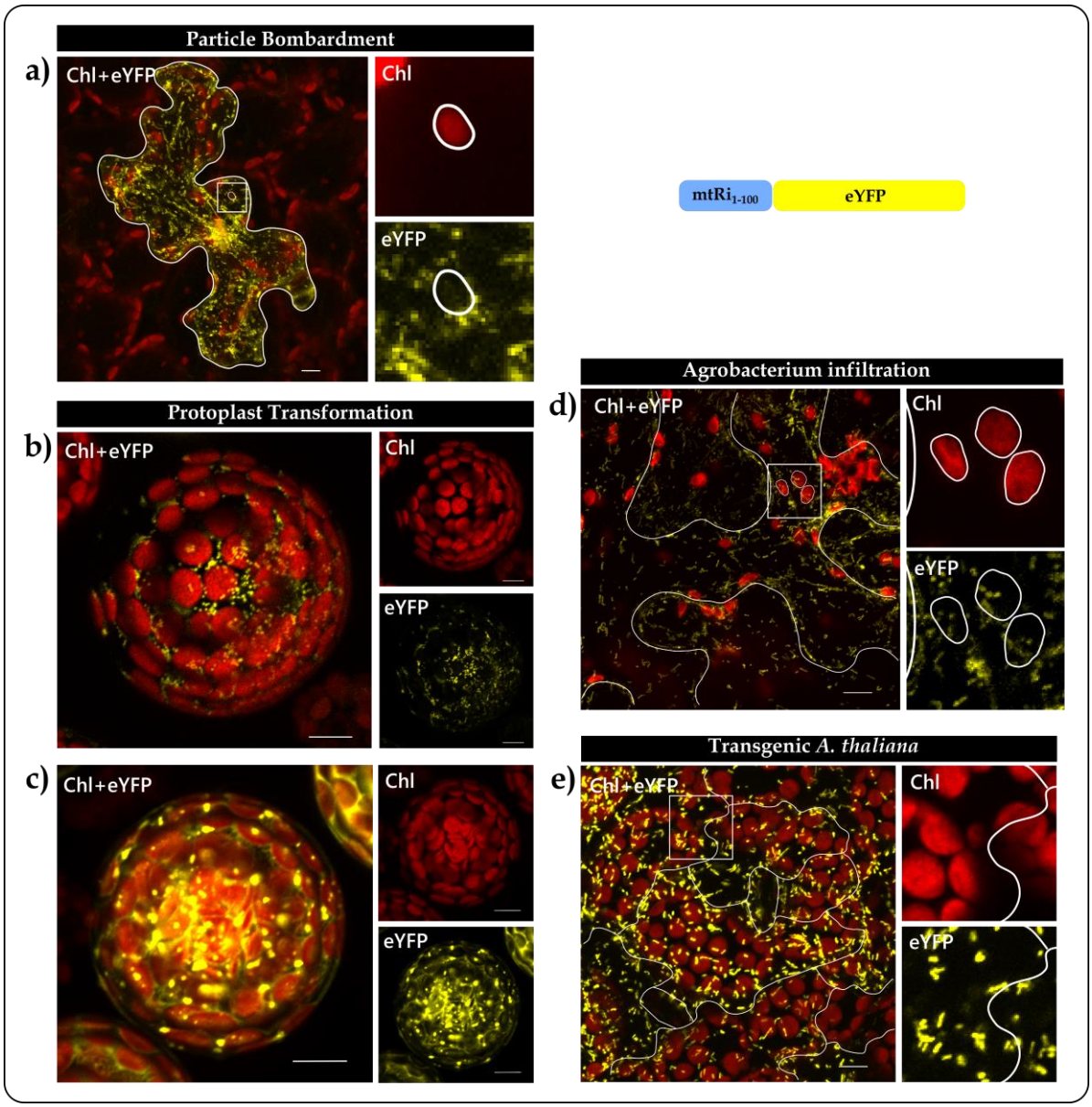


Figure 2.2 Subcellular localization of the candidate protein mtRi$_{1\text{-}100}$/eYFP. The subcellular protein localization of mtRi$_{1\text{-}100}$/eYFP was determined by confocal laser scanning microscopy after particle bombardment of leaf epidermal cells of *Arabidopsis thaliana* (panels **a**), protoplast transformation of *Arabidopsis thaliana* (panels **b** and **c**), *Agrobacterium* infiltration of *Nicotiana benthamiana* (panels **d**), and in transgenic *Arabidopsis thaliana* lines (panels **e**). In all instances, mtRi$_{1\text{-}100}$/eGFP is localized to mitochondria. At high expression levels, mtRi$_{1\text{-}55}$/eYFP often accumulates as aggregates in transformed protoplast (panel **c**). The scale bars correspond to 10 µm. (See legend of figure 2.1 for more details)

(b) Glycine cleavage system subunit-H 1 (GCS-H1)

Among the four candidate proteins analyzed, GCS-H1 is addressed as a mitochondrial protein in literature (Douce et al., 2001). It is a part of a large glycine cleavage system (GCS) comprised of four protein sub-units namely P-protein, H-protein, T-Protein and L-protein (Oliver et al., 1990). In mitochondria, GCS catalyses photorespiratory release of CO_2 and NH_3 after decarboxylation of glycine. The H subunit of GCS acts as a mobile substrate that commutes between the other

subunits (Douce et al., 2001). Three isogenes of GCS-H are encoded in *Arabidopsis thaliana* genome namely *GCS-H1*, *GCS-H2* and *GCS-H3* (Bauwe and Kolukisaoglu 2003) and the transcripts of *GCS-H1* are most abundant in the cell. The subcellular localization of GCS-H1 protein (further referred as **GCS**) was analysed in this study.

The gene sequence encoding N-terminal 100 amino acids of GCS fused to eYFP (GCS_{1-100}/eYFP) were expressed in plant cells via above experimental approaches. In all four assays, GCS_{1-100}/eYFP showed a clear dual localization to mitochondria and plastids (Figure 2.3). Comparable fluorescence signals were observed in two organelles each with particle bombardment, protoplast transformation and with *Agrobacterium* infiltration. While in case of stable transgenic lines, fluorescence signal intensity was comparatively stronger in mitochondria indicating its preferential accumulation in this organelle (Figure 2.3d).

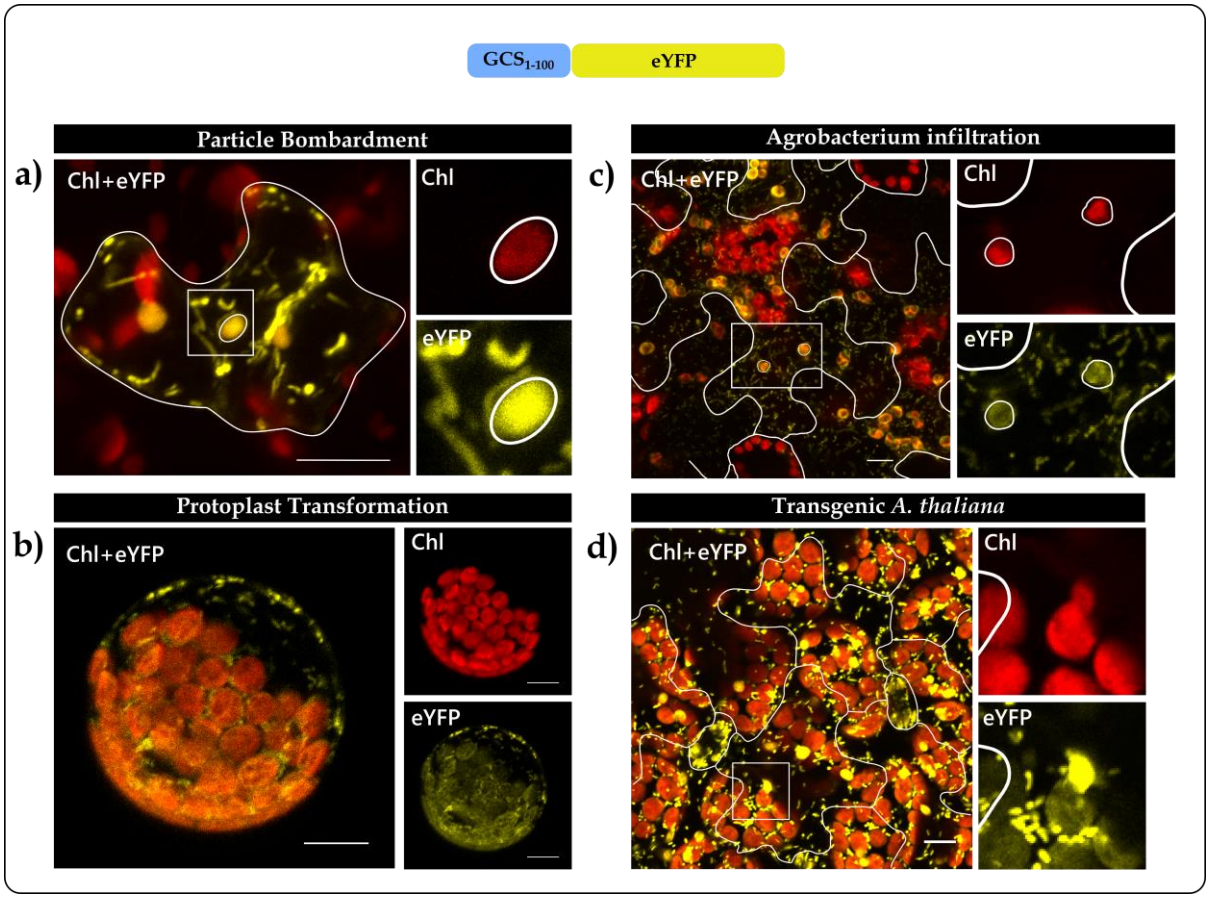


Figure 2.3 Subcellular localization of the candidate protein GCS_{1-100}/eYFP. The subcellular protein localization of GCS_{1-100}/eYFP was determined by confocal laser scanning microscopy after particle bombardment of leaf epidermal cells of *Arabidopsis thaliana* (panel **a**), protoplast transformation of *Arabidopsis thaliana* (panel **b**), *Agrobacterium* infiltration of *Nicotiana benthamiana* **(**panel **c**), and in transgenic *Arabidopsis thaliana* lines (panel **d**). In all instances, GCS_{1-100}/eYFP is localized to both mitochondria and plastids. Representative cells showing dual localization of the candidate proteins in both mitochondria and chloroplasts are presented as overlay images of the chlorophyll channel (displayed in red) and the eYFP channel (displayed in yellow). The scale bars correspond to 10 µm. (For further details see the legend of Figure 2.1)

(c) Co-chaperone/nucleotide exchange factor mitochondrial GrpE I (Mge1 or GrpE)

Second candidate protein in this comparative analysis was MgeI, which has been described as a monospecific mitochondrial protein in the literature (Hu et al., 2012). MgeI belongs to a highly conserved and ubiquitous GrpE (GroP like gene E) family of proteins; function as nucleotide exchange factor in Hsp70/DnaK chaperone machinery. The mitochondrial GrpEs are associated with inner mitochondrial membrane translocase complex and assist Hsp70 in pre-protein translocation and maturation (Bolliger et al., 1994). In model plant *Arabidopsis thaliana* four organelle targeted homologues of GrpE namely mitochondrial GrpE 1, 2 (*Mge1*, *Mge2*) and chloroplastic GrpE1, 2 (*Cge1*, *Cge2*) are encoded in the cell nucleus (Hu et al., 2012). The subcellular localization of Mge1 (further referred as **GrpE** in this study) was examined via above-mentioned experimental approaches.

Surprisingly, different experimental approaches yielded variable subcellular localization pattern of $GrpE_{1-100}$/eYFP in transformed cells. When analyzed via particle bombardment, $GrpE_{1-100}$/eYFP was found dually targeted in approximately 50% of the analyzed cells, but monospecifically targeted to mitochondria in others within the very same experiment (Figure 2.4a-b). Second transient assay involving mesophyll protoplast transformation also indicated dual localization of chimeric fluorescent protein still, with preferential accumulation in mitochondria (Figure 2.4d). In case of *Agrobacterium* infiltration assays, not all cells were homogeneously transformed and patches of high and low fluorescence intensity were visible in the transformed leaf tissue. The low expression of the $GrpE_{1-100}$/eYFP constructs leads to apparent monospecific targeting to mitochondria, while cells with high fluorescence intensity showed clear dual targeting to plastids and mitochondria (Figure 2.4c). In the latter case, there was a clear correlation of expression rate and targeting behavior, which can be studied in such assays in a single step.

Contrastingly, in the stable transgenic lines, $GrpE_{1-100}$/eYFP accumulated exclusively in the mitochondria of the analysed leaf tissue (Figure 2.4e). It could be assumed that the post-transport processes, like protein turnover rates, rather than incompetence of the transit peptide to mediate plastid import are responsible for the observed lack of $GrpE_{1-100}$/eYFP accumulation in chloroplasts of transgenic lines.

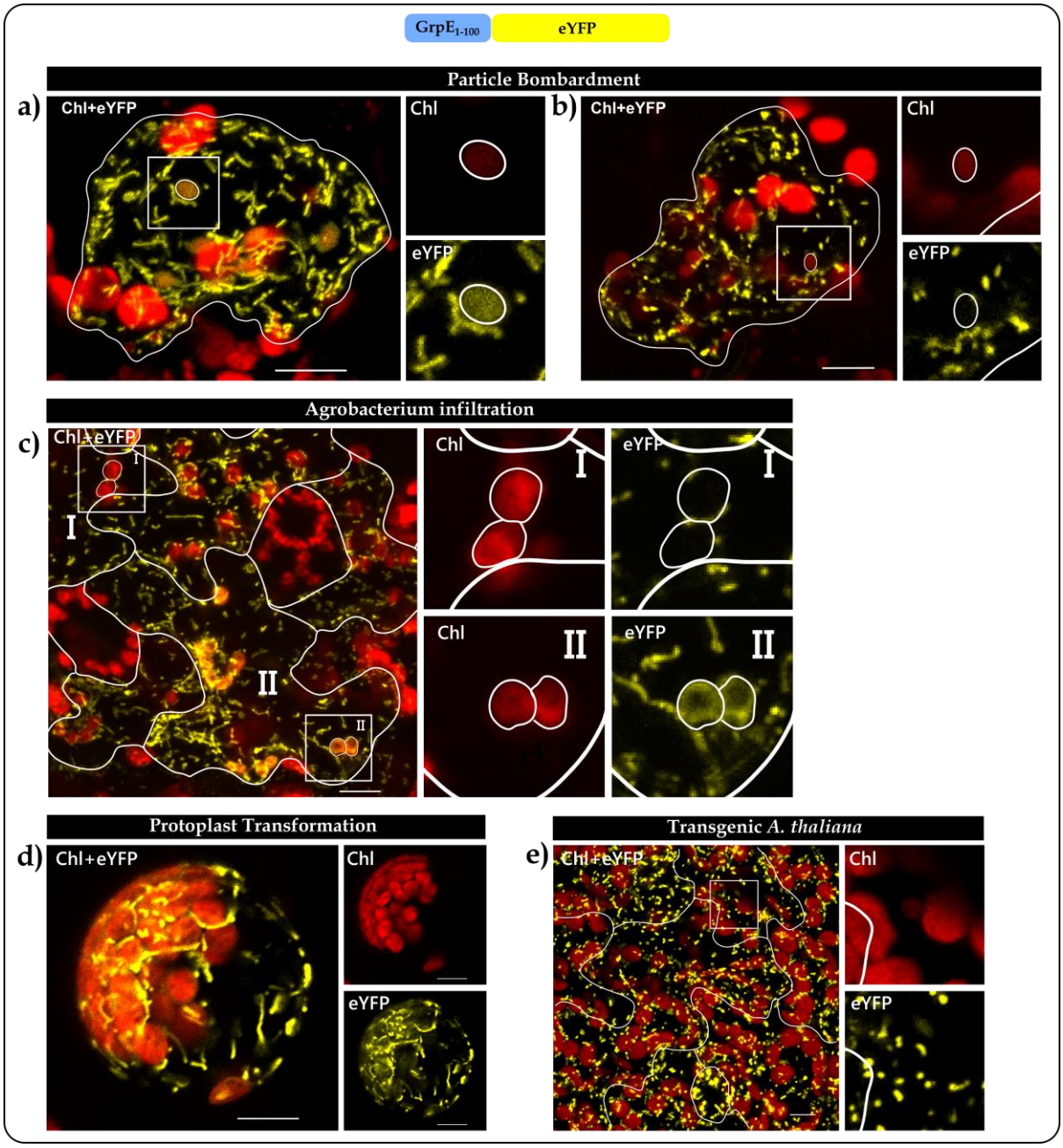


Figure 2.4 Variable subcellular localization of the candidate protein $GrpE_{1-100}$/eYFP. The subcellular protein localization of $GrpE_{1-100}$/eYFP was determined by confocal laser scanning microscopy after particle bombardment of leaf epidermal cells of *Arabidopsis thaliana* (panels **a** and **b**), *Agrobacterium* infiltration of Nicotiana benthamiana (panels **c**), protoplast transformation of *Arabidopsis thaliana* (panels **d**), and in transgenic *Arabidopsis thaliana* lines (panels **e**). The particle bombardment can lead to either dual localization of the candidate protein in both mitochondria and chloroplasts (**a**) or to mitochondrial localization solely (**b**). In case of *Agrobacterium* infiltration (**c**) the dual targeting of the reporter protein was observed only in cells with strong expression of the candidate constructs (cII), not in those with low expression levels (cI), which show solely mitochondrial accumulation. The transformed protoplast showed dual localization (**d**) while, exclusive mitochondria localization was observed in case of stable transgenic lines (**e**). Representative cells are displayed as overlay images of the chlorophyll channel (displayed in *red*) and the eYFP channel (displayed in *yellow*). The scale bars correspond to 10 µm. (For further details see the legend of Figure 2.1)

(d) Mitochondrial elongation factor Tu (EF-Tu)

Third protein chosen for this study was mitochondrial EF-Tu from *Arabidopsis thaliana*. Elongation factor Tu (EF-Tu) is an essential component of protein translation machinery and belongs to a large family of GTP-binding proteins (Krab and Parmeggiani, 1998). During protein synthesis, EF-Tu facilitates transfer of aminoacylated t-RNA to the A-site of ribosome with concomitant hydrolysis of GTP. Besides their role in protein synthesis, EF-Tu play an important role in protein folding and a variety of cellular functions (reviewed by Fu et al., 2012). The genome of *Arabidopsis thaliana* encodes two organelle targeted EF-Tu namely plastidial EF-Tu and mitochondrial EF-Tu that are phylogenetically related to cyanobacterial EF-Tu and eubacterial Ef-Tu respectively (Kuhlman and Palmer 1995).

In line with the results obtained for $GrpE_{1\text{-}100}$/eYFP, the particle bombardment of $EF\text{-}Tu_{1\text{-}100}$/eYFP also indicated its preferential targeting into mitochondria but only in few instances, a faint fluorescence signal was detectable in plastids (Figure 2.5a-b). Moreover, even these signals were difficult to visualize due to the considerable background resulting from the rapid movement of mitochondria during image acquisition. This was particularly well visible when single image planes were compared with a complete Z-stack projection of the entire cell (Figure 2.5b). Similarly, in protoplast transformation assays, $EF\text{-}Tu_{1\text{-}100}$/eYFP showed a stronger fluorescence signals in mitochondria while the signals in plastids were rather vague (Figure 2.5d). Variable protein localization, depending on the expression level of construct, was observed in *Agrobacterium* infiltrated cell (Figure 2.5c). While, in stable transgenic plants $Ef\text{-}Tu_{1\text{-}100}$/eYFP accumulated solely in mitochondria (Figure 2.5e).

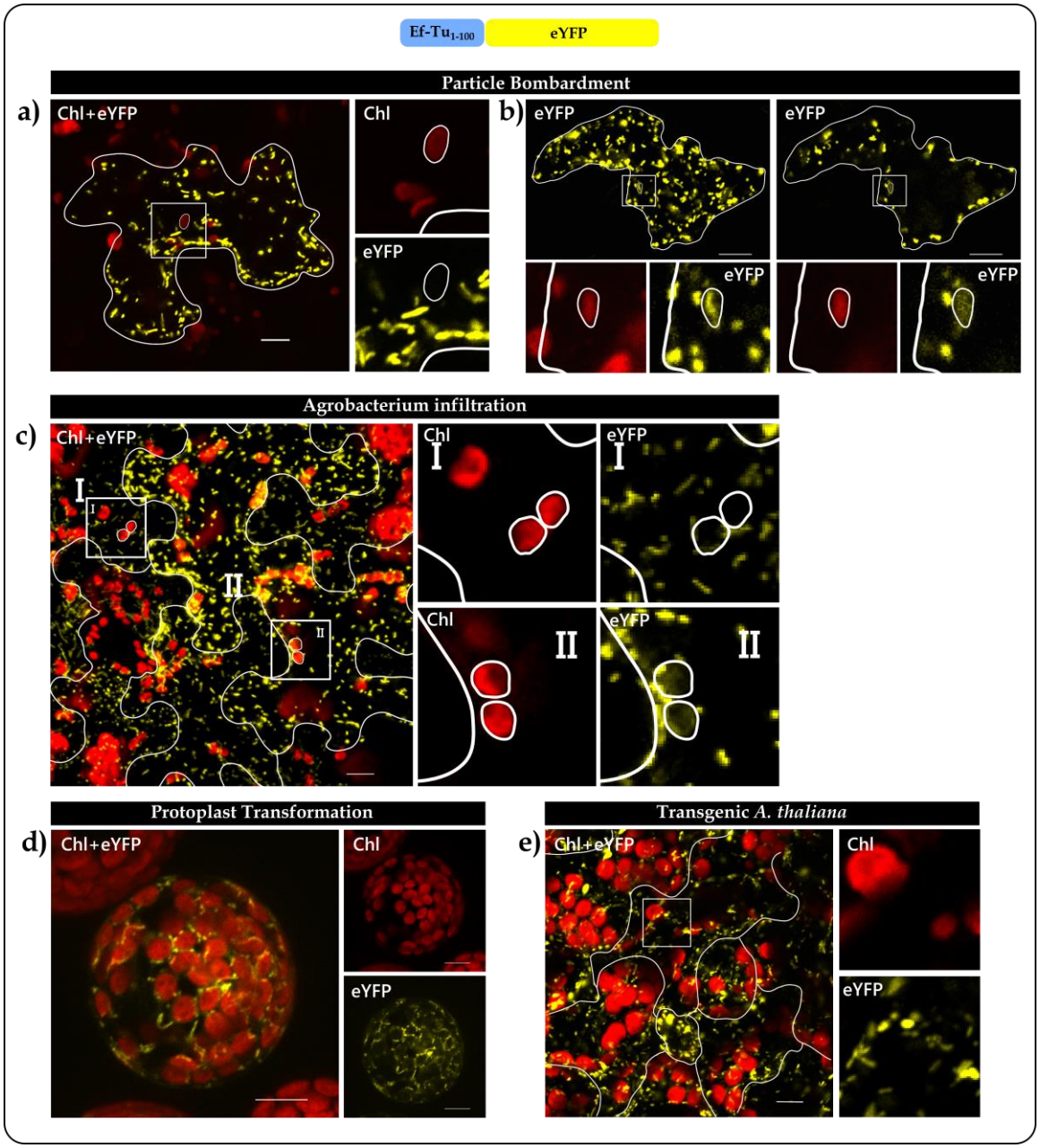


Figure 2.5 Subcellular localization of the candidate protein Ef-Tu$_{1-100}$/eYFP. The subcellular protein localization of Ef-Tu$_{1-100}$/eYFP was determined by confocal laser scanning microscopy after particle bombardment of leaf epidermal cells of *Arabidopsis thaliana* (panels **a** and **b**), *Agrobacterium* infiltration of *Nicotiana benthamiana* (panels **c**), protoplast transformation of *Arabidopsis thaliana* (panels d), and in transgenic *Arabidopsis thaliana* lines (panels **e**). The particle bombardment lead to either a mitochondrial localization (**a**) or dual localization with preferential accumulation to mitochondria (**b**). In latter case the plastid localization is often masked by high fluorescence signals coming from mitochondria as depicted by a Z-stack of several images (panel **b** *left*) or a single plane image of transformed cell (panel **b** *right*). In case of *Agrobacterium* infiltration (**c**) the dual targeting of the reporter protein was observed only in cells with strong expression of the candidate constructs (cII), not in those with low expression levels (cI), which show solely mitochondrial accumulation. The transformed protoplast showed dual localization with preferential accumulation in mitochondria (**d**), while exclusive mitochondria localization was observed in case of stable transgenic line (**e**). Representative cells are displayed as overlay images of the chlorophyll channel (displayed in *red*) and the eYFP channel (displayed in *yellow*). The scale bars correspond to 10 µm. (for further details see the legend of Figure 2.1)

(e) Peptide deformylase 1 B (PDF-1B or PDF)

The last candidate protein in this analysis was a protein from peptide deformylase (PDF) superfamily. Peptide deformylases facilitate the removal of formyl group from N-terminal methionine of newly synthesized proteins in endoymbiotic organelles (Serero et al., 2001). In *Arabidopsis thaliana*, two organelle targeted peptide deformylases, namely PDF1A (DEF1) and PDF1B (DEF2), are encoded in the nucleus. In a subcellular localization study by Giglione et al. (2000), PDF1A was shown to be targeted to mitochondria; while PDF1B was dually targeted to plastids and mitochondria. A separate study by Dinkins et al. (2003) demonstrated that PDF1A is a dually targeted protein and that dual localization of PDF1A is a consequence of alternative translation start at a different AUG codon. Here, the targeting specificity of PDF1B (further referred as **PDF**) was examined.

In all four assays, $PDF_{1\text{-}100}$/eYFP showed a clear dual localization to both mitochondria and plastids and no significant variability was observed within different experimental approaches (Figure 2.6). Still, in transient transformation assays, the fluorescence signal intensity was less pronounced in mitochondria than in chloroplasts. Similarly, in case of stable transgenic lines, $PDF_{1\text{-}100}$/eYFP showed stronger accumulation in plastids, which again indicates the preference of dually targeted proteins for one or the other organelle (Figure 2.6d).

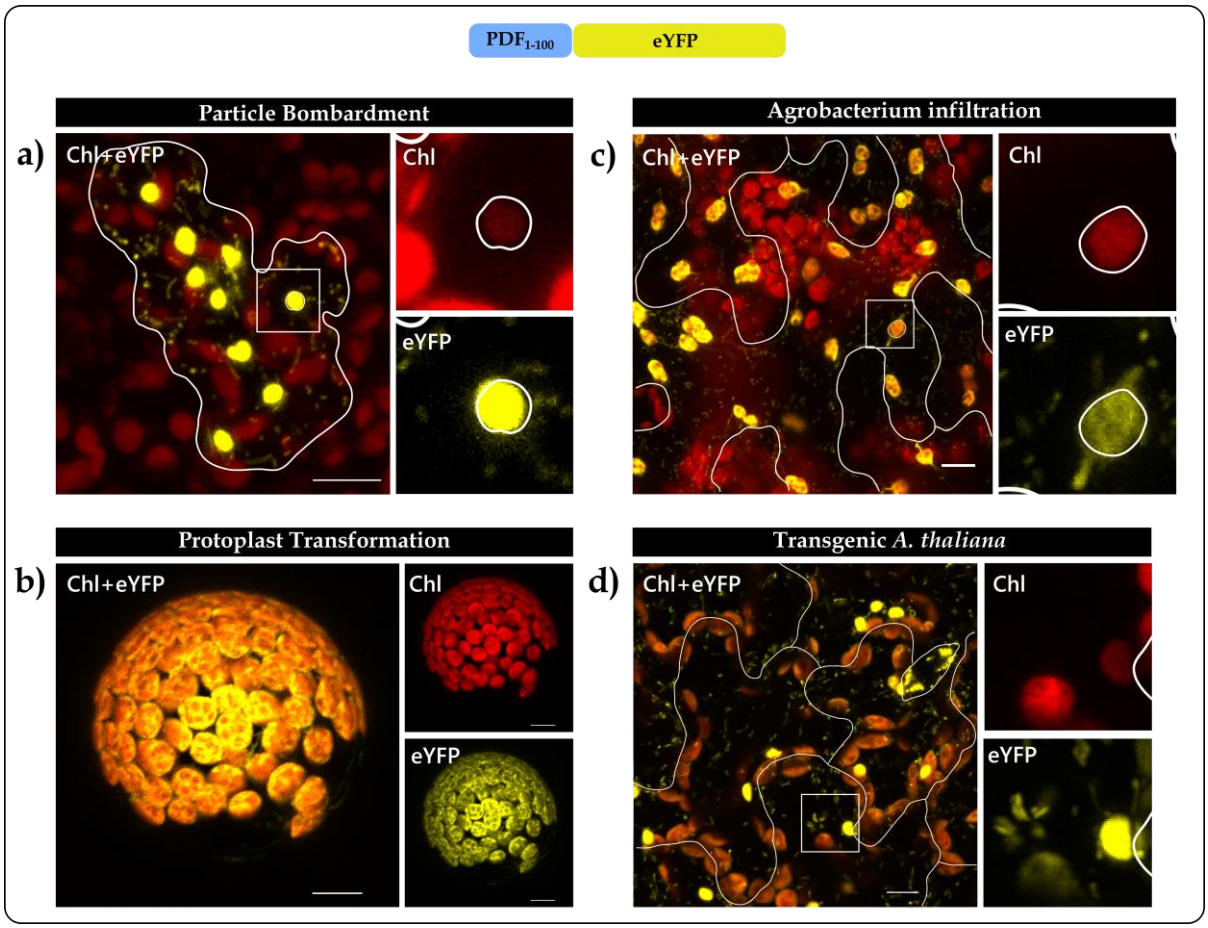


Figure 2.6 Subcellular localization of the candidate protein PDF_{1-100}/eYFP. The subcellular protein localization of PDF_{1-100}/eYFP was determined by confocal laser scanning microscopy after particle bombardment of leaf epidermal cells of *Arabidopsis thaliana* (panels a), protoplast transformation of *Arabidopsis thaliana* (panels b and c), *Agrobacterium* infiltration of *Nicotiana benthamiana* (panels d), and in transgenic *Arabidopsis thaliana* lines (panels e). In all instances, PDF1-100/eYFP showed dual localization in both plastids and mitochondria with preferential accumulation in plastids. Representative cells showing dual localization of the candidate proteins in both mitochondria and chloroplasts are presented as overlay images of the chlorophyll channel (displayed in *red*) and the eYFP channel (displayed in *yellow*). The scale bars correspond to 10 µm. (For further details see the legend of Figure 2.1)

2.3 Discussion

One remarkable outcome of the comparative evaluation of largely similar experimental systems utilized to determine the organellar targeting of nuclearly encoded proteins in plants is that divergent or even contradictory results can still all be valid and true (summarized in Table 2.1). This appears impossible at first glance but may be explained by the fact that neither of the approaches reflects the organelle transport process in an unbiased manner. Instead, they all have their particular drawbacks and strengths, which become evident only if they are directly compared with each other.

2.3.1 Transient vs. Stable Transformation

A good example for such variability is $GrpE_{1-100}$/eYFP, which shows dual targeting in all transient assays described here but accumulates in transgenic plants solely in mitochondria (Figure 2.4a-b). At first glance, this might suggest that all transient assays lead to artificial chloroplast import of the protein. However, this is probably a too simplistic view. Instead, the discrepancy observed after transient vs. stable transformation is probably a consequence of the different time-scales at which the analyses are performed. In transient assays, subcellular localization of the reporter construct is determined by fluorescence microscopy usually within 16–72 h after transformation, i.e., the protein is present in the cell for only a limited time period before being analyzed. In contrast, even if young tissue of transgenic plants is used for microscopy, the cells have expressed the reporter gene for several days or even weeks prior to analysis. Thus, regulatory processes, like protein turnover or counter selection caused by incompatibility of the reporter construct with the cell metabolism, can exert a major effect on the accumulation of the protein in the organelles. Regulating the time-point of transgene expression, for example by using inducible promoters, could circumvent such effects but this is not a common practice while studying protein targeting specificity.

Table 2.1 Localization of candidate proteins obtained with different experimental approaches.

	Particle Bombardment (*Arabidopsis*)	**Protoplast Transformation (*Arabidopsis*)**	**Agrobacterium Infiltration (*Nicotiana*)**	**Transgenic Plants (*Arabidopsis*)**	**Mass Spectrometry** (*Arabidopsis*)**	**Literature Data**
GrpE/eYFP (At5g55200*)	a) Dual b) Mito	**a) Dual**	a) Dual b) Mito	**a) Mito**	**Mito**	Dual (Onion)[1] Mito (Arabidopsis)[1] Dual (Pea)[2]
EF-Tu/eYFP (At4g02930*)	**a) Mito** b) Dual	a) Dual b) Mito c) Aggregates	a) Dual b) Mito	**a) Mito**	**Mito** Plastid	Dual (Pea)[2]
GCS/eYFP (At2g35370*)	**a) Dual**	**a) Dual** b) Aggregates	**a) Dual**	**a) Dual**	**Mito** **Plastid**	Dual (Pea)[2]
PDF/eYFP (At5g14660*)	**a) Dual**	**a) Dual**	**a) Dual**	**a) Dual**	**Plastid**	Dual (Pea)[2] Dual (Onion)[3]
mtRi/eYFP (GB:X79332.1*)	**a) Mito** b) Cytosol + Nucleus	a) Mito b) Cytosol + Aggregates	a) Mito b) Cytosol	**a) Mito**	-	Mito (Pea)[2]
FNR/eGFP (GB:M86349.1*)	**a) Plastid**	a) Plastid b) Cytosol + Nucleus	a) Plastid	**a) Plastid**	-	Plastid (Pea)[2]

Dual: localization in mitochondria and plastids; *Mito/Plastid*: localization exclusively in mitochondria/plastids; Aggregates: protein aggresomes in the transformed cell. a, b and c address distinct localization in different cells of the same experiment. Preferential or exclusive localization is highlighted in **bold**.
*Accession number of the corresponding candidate gene; **Data obtained from MASCP GATOR (Joshi et al., 2011; Mann et al., 2013) and SUBA4 Databases (Hooper et al., 2017); [1]Van Aken et al., 2009; [2]Baudisch et al., 2014; [3]Giglione et al., 2000

A second point to be considered is the position at which the *T-DNA* carrying the candidate gene is inserted into the nuclear genome of the stably transformed plant line. In most instances, this insertion will take place in a non-essential region of the genome but it cannot be ruled out that occasionally also a gene is affected that plays a role in the subcellular targeting or accumulation of proteins. Consequently, it is essential to analyze more than a single transgenic line to prevent potential misinterpretation.

2.3.2 Protoplasts vs. Intact Tissue

Even the various transient transformation systems have each their specific peculiarities. In transient transformation assays performed with intact tissue, like particle bombardment, single cells of known and defined origin are transformed. In contrast, isolated protoplasts always consist of a mixture of differentiated cell types. For example, protoplasts prepared from leaves of dicotyledonous plants will comprise not only mesophyll cells from palisade and spongy parenchyma but also epidermal and stomatal cells. Taking into account that protoplasts from different plant tissues can show differences in their protein targeting characteristics (Faraco et al., 2011), it can be assumed that also unequally differentiated cells in the same assay will show different transport properties.

Furthermore, the procedure of protoplast isolation and transformation is stressful for the cell. It has been reported that certain stress conditions can even lead to the release of proteins from organelles (for example plastids) resulting in their accumulation in the cytosol (Kwon et al., 2013). Nevertheless, maybe even more important is the fact that stress usually induces the expression of genes encoding chaperones like Hsp70 (Wang et al., 2004), which are involved also in protein import into mitochondria and chloroplasts (Zhang and Glaser, 2002). Together, these findings might explain why organelles can have different import characteristics in transient assays comparing protoplasts with cells of intact tissue.

2.3.3 Presence or Absence of *Agrobacterium* in the Assays

Agrobacterium infiltration of *N. benthamiana* usually yields a large number of transformed cells within an intact tissue that can be easily identified and analyzed at the subcellular level. This method therefore appears particularly suitable for large-scale protein localization studies. However, in addition to the fact that proteins analyzed are often of heterologous origin, the results may be confounded by symptoms resulting from the inoculation of the plant tissue with bacterial cells. In fact, it was shown that in leaf areas infiltrated with *A. tumefaciens*, defense reactions and chlorosis are sometimes induced (Pruss et al., 2008). Furthermore, depending on the *Agrobacterium* strain used for transformation, altered phytohormone levels in the plant tissue have been described (Erickson et al., 2014). At this point, it cannot be ruled out that such stress-related plant responses can well have an influence also on the organelle targeting of proteins. In fact, it is worthwhile to systematically analyze if and how such targeting processes depend on physiological conditions like the energy load of the cell or its redox status.

2.3.4 Choice of the Experimental System for Protein Transport Studies

Considering the pros and cons of all experimental approaches described here, it becomes clear that there is no perfect method to study the specificity of protein targeting into organelles of intact plant cells. While the choice of the experimental system appears almost negligible in those cases in which a given candidate protein shows efficient transport with comparable rates into both endosymbiotic organelles (e.g., GCS), it is much more important if the protein shows preferential targeting to one or the other organelle (e.g., EF-Tu). In this case, more than a single approach is required to avoid misinterpretation, although it is obviously not a serious option to demand for all assays when analyzing a candidate protein.

Remarkably, even biochemical or proteomics approaches, which are often considered to be unbiased, have their specific inherent deficiencies. In principle, they depend strictly on the quality and, in particular, purity of the organelles studied. However, such extremely pure organelles cannot usually be obtained with standard isolation procedures and a certain degree of cross-

contamination of, for example, mitochondria with proplastids is almost impossible to prevent (e.g., Keech et al., 2005; Rödiger et al., 2010). Furthermore, the methods of organelle isolation were usually established for “typical” organelles (e.g., mesophyll chloroplasts) and thus do not necessarily apply to other subtypes (e.g., epidermis chloroplasts) which can have deviating physico-chemical characteristics. Therefore, the data sets of organelle proteome analyses might well miss some organellar proteins but, on the other hand, inevitably comprise also a number of proteins from contaminating organelles. To cope with that, threshold levels are usually implemented to separate “true” from “wrong” results. However, such threshold levels are arbitrary and it might well be that a highly abundant contaminating protein shows an even higher value than an actual organelle protein of low abundance. This problem is particularly evident in the case of dually targeted proteins with preferential localization in only one of the two organelles and might be the reason for the sometimes discrepant proteomics data, e.g., for EF-Tu, which was found by mass spectroscopy always in mitochondria but only in a single case also in plastids (Helm et al., 2014; see also SUBA4 database).

In conclusion, the experimental system should be carefully chosen depending on the question to be answered, since each method addresses different aspects of the transport process. For example, if tissue specificity of the targeting process is anticipated, transgenic plants are the only valuable option. If instead targeting specificity in a homologous system needs to be studied, particle bombardment and/or protoplast transformation are the methods of choice. The latter cannot be applied though if stress needs to be avoided, which holds true also for *Agrobacterium* infiltration. In this case, particle bombardment or, even better, transgenic plants have to be utilized. Finally, if the principal property of a candidate protein to interact with the import machinery of an organelle is of interest, *in organello* import experiments performed with isolated intact organelles are still an option.

Chapter-3
The self-assembling split-fluorescent protein system to determine targeting specificity of nuclearly encoded organelle proteins

3.1 Background

A comparative analysis of established fluorescence microscopy based approaches highlighted that the 'potential' dual targeting specificity of organelle proteins might remain undetected especially in those instances where the targeting of these proteins is preferential towards one of the organelles. For example determining dual targeting specificity of GrpE and Ef-Tu was particularly tedious and error-prone with the conventional FP-tagging approaches (Chapter 2). Besides this, the ectopic expression of fluorescent protein fusions commonly results in considerable background signals and the large size and robust folding of the reporter protein may perturb the membrane transport of candidate proteins.

One possibility to circumvent these problems could be to detect the fluorescence signals, coming from each organelle, in separate cells. This has become technically feasible by the invention of the self-assembling green fluorescent protein (*sas*plit-GFP) technology, a popular approach developed recently to determine targeting specificity of proteins *in vivo* (Cabantous et al., 2005). Spontaneous self-assembly of split-GFP relies on two highly engineered fragments derived from the 'superfolder' GFP variant, sfGFP. The large fragment, GFP1-10 OPT (further referred to as GFP1-10), comprises ten N-terminal antiparallel β-sheets of GFP. The smaller fragment, GFP11 M3 (further referred to as GFP11), is only 16 AA long and represents the C-terminal 11th β-sheet of GFP. When brought into close proximity, the two fragments can assemble spontaneously to reconstitute the functional fluorophore without the need of an additional interacting partner (Figure 3.1a) (Cabantous et al., 2005). For subcellular protein localization studies, GFP1-10 is fused to a transport signal of known organelle specificity and analysed together with a chimeric protein comprising the candidate protein and GFP11. Fluorescence complementation is achieved specifically and exclusively if the transport signal of the candidate mediates transport of GFP11 into the organelle housing GFP1-10. The fluorescence signal remains limited to the compartment

containing GFP1-10, irrespective of whether the GFP11-fused candidate protein is targeted to further subcellular compartments or not (Figure 3.1b). Thus, the *sa*split-GFP technology allows selective *in vivo* imaging of a protein of interest in a respective compartment with enhanced signal-to-noise ratio. Furthermore, the small GFP11 might have a lower propensity for interfering with membrane transport in comparison to the full-length fluorescent proteins used for direct FP-tagging. Smaller size tags were previously shown to be advantageous over large FP-tags for live cell imaging (Andresen et al., 2004; Giepmans et al., 2006).

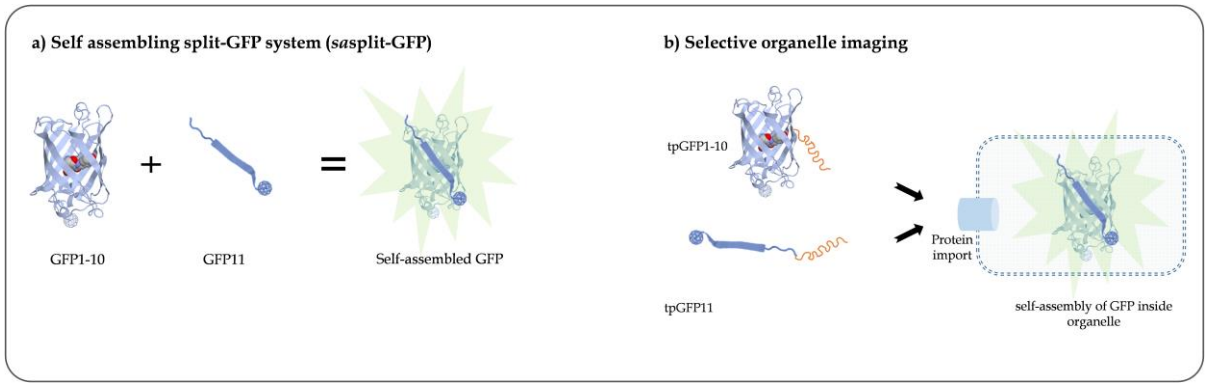


Figure 3.1 Schematic drawing illustrating the principle of the *sa*split-GFP system. **(a)** Two non-fluorescing fragments, GFP1-10 and GFP11, can self-assemble to generate a fluorescing GFP molecule. **(b)** Transport of both GFP chimeras via organelle specific transport signals into the same organelle is essential to achieve fluorescence signals.

The *sa*split-GFP system has already been applied to a wide range of organisms and adapted to elucidate a variety of cellular functions or processes including, for example, *in vivo* protein solubility, subcellular localization of pathogen effectors, endogenous protein labelling for *in vivo* imaging, protein-protein interaction studies, and membrane protein topology determination (Cabantous and Waldo 2006; Van Engelberg and Palmer, 2010; Machettira et al., 2011; Cabantous et al., 2013; Kamiyama et al., 2016; Henry et al., 2017). The principal suitability of the *sa*split-GFP system for *in vivo* imaging of protein targeting in plant cells has also been demonstrated recently (Park et al., 2017). In this case, transgenic *Arabidopsis thaliana* lines expressing the organelle-targeted GFP1-10 receptor were transiently transformed with constructs expressing a candidate protein fused to GFP11 tag. However, the requirement of transgenic 'receptor' plant lines and low transient transformation efficiency hampers the utilization of this approach for high throughput protein targeting studies. Along with a limited yield of transformed cells, the relatively low brightness of the *sa*split-GFP prevented the visualization of proteins targeted in low amounts into the organelles.

Therefore, here the *sa*split-GFP system was optimized for the analysis of protein targeting specificity in plant cells and the effect of multimerization of the GFP11 tag, on the intensity of the fluorescence signals inside mitochondria and plastids, was assessed. In addition, a *Golden*

Gate-based vector toolkit named PlaMiNGo (Analysis of Plastid and/or Mitochondrial targeted proteins N- terminally fused to GFP11 tags via *Golden Gate* cloning) was developed to facilitate high-throughput analyses of candidate proteins with high transformation efficiency and enhanced signal-to-noise ratio. The targeting specificity of previously characterized dually targeted proteins was "re-evaluated" with this approach and such dual targeting was newly detected for several proteins that were previously characterized as being targeted to a single organelle only. Importantly, plastid targeting of these proteins was independently confirmed using a phenotype complementation-based approach in stable transgenic *Arabidopsis* plants, to confirm the *in vivo* relevance of results obtained with the PlaMiNGo system.

3.2 Results

3.2.1 The *sa*split-GFP system to determine protein targeting specificity

In order to study protein targeting into plastids, the transit peptide of a chloroplast protein, ferredoxin-NADP$^+$-oxidoreductase of spinach (FNR_{1-55}; Zhang et al., 2001), was fused to the GFP1-10 receptor (the large fragment of the *sa*split-GFP system) to facilitate its localization into plastids. For mitochondrial localization of GFP1-10, the N-terminal 100 amino acid residues comprising the presequence of the mitochondrial Rieske Fe/S protein of potato ($mtRi_{1-100}$; Emmermann et al., 1994) were used. Both transport signals had previously been characterized for their targeting specificity to a single organelle with *in vivo* and *in vitro* approaches (section 2.2.3 and Rödiger et al., 2011). For the initial experiments evaluating the suitability of the *sa*split-GFP system for our purposes, each transport signal was likewise combined with the $GFP11_{x7}$ tag (small fragment of *sa*split-GFP system). A seven-fold repeat of this GFP11 tag ($GFP11_{x7}$; separated via a five amino acid linker) was used, as such multiple GFP11 tags had been reported to intensify the fluorescence signals in mammalian cells (Kamiyama et al., 2016). When these constructs were co-infiltrated in *N. benthamiana* leaf epidermal cells with the FNR_{1-55}/GFP1-10 and $mtRi_{1-100}$/GFP1-10 gene constructs, fluorescence signals were exclusively obtained in those instances in which the same transport signal was present in both chimeras, i.e. in plastids after co-expression of FNR_{1-55}/GFP1-10 and FNR_{1-55}/$GFP11_{x7}$ and in mitochondria after co-expression of $mtRi_{1-100}$/GFP1-10 and $mtRi_{1-100}$/$GFP11_{x7}$ (Figure 3.2). When infiltrated alone or in reciprocal combinations, neither of these constructs generated any detectable fluorescence signal (Figure 3.3). This demonstrates the suitability and specificity of the *sa*split-GFP system for the analysis of protein targeting.

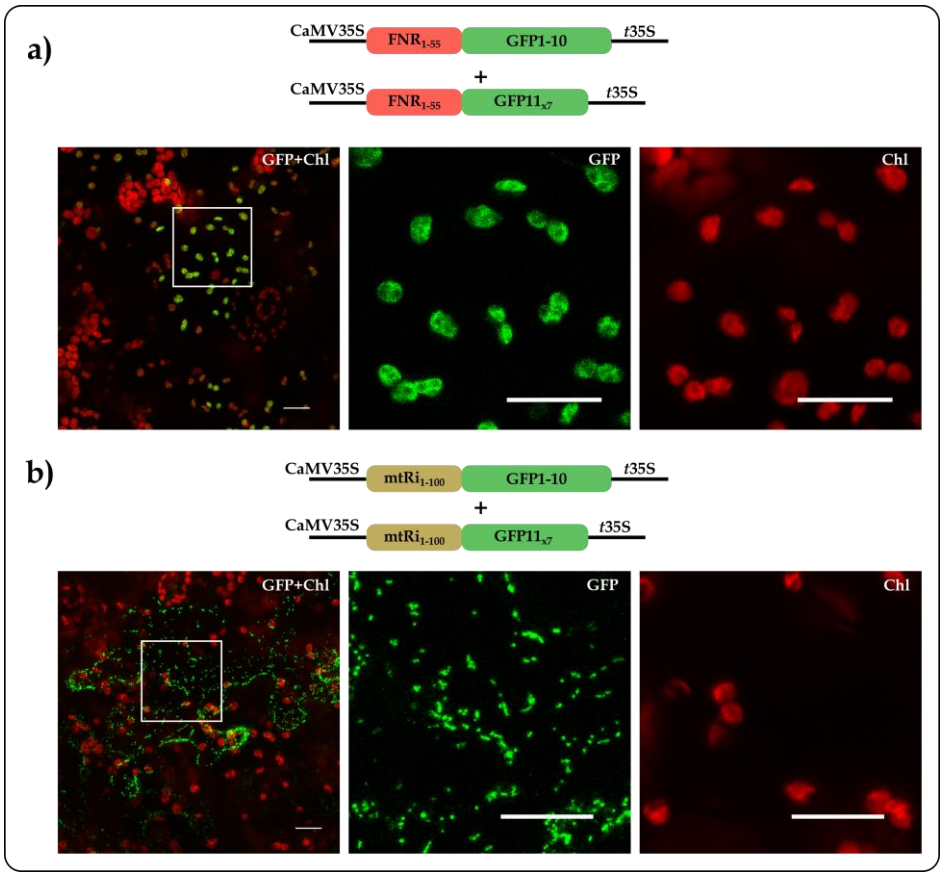


Figure 3.2 Establishment of the *sa*split-GFP system for *in vivo* organelle imaging. The coding sequences of **(a)** FNR_{1-55}/GFP1-10 and FNR_{1-55}/$GFP11_{x7}$ **(b)** $mtRi_{1-100}$/GFP1-10 and $mtRi_{1-100}$/$GFP11_{x7}$ were transiently co-expressed after *Agrobacterium* co-infiltration into the lower epidermis of *Nicotiana benthamiana* leaves and analyzed by confocal laser scanning microscopy. Image acquisition of transformed cells was done with 20-x objective in several Z-stacks, which were subsequently stacked for maximum intensity projection. Representative cells (*left panels*) are presented as overlay images of the chlorophyll channel (displayed in red) and the GFP channel (displayed in green). The strong chlorophyll signals in the background are derived from the larger chloroplasts of untransformed mesophyll cells underneath the epidermal cell layers. The *squares* highlight areas of the transformed cells that are shown in higher magnification separately for the chlorophyll channel (*middle panels*) and the GFP channel (*right panel*) as indicated. Scale bars correspond to 20 µm.

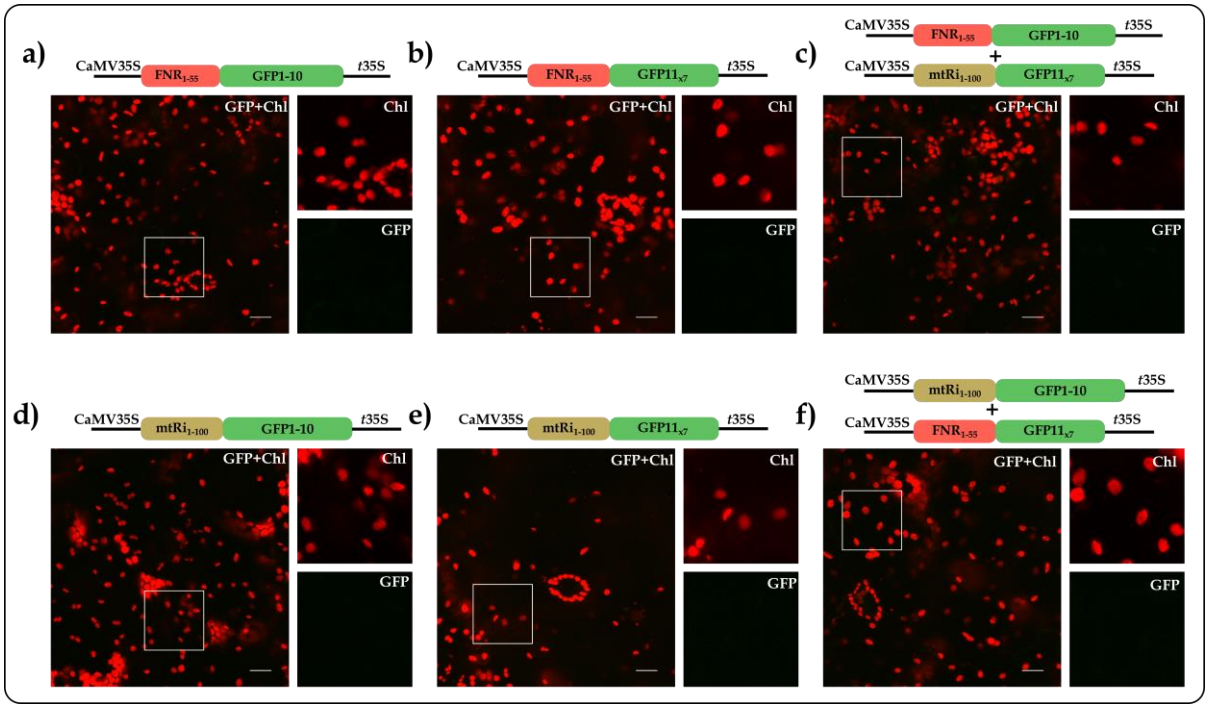


Figure 3.3 Experimental controls for *sa*split-GFP system. The coding sequences **(a)** FNR_{1-55}/GFP1-10, **(b)** FNR_{1-55}/$GFP11_{x7}$, **(d)** $mtRi_{1-100}$/GFP1-10 or **(e)** $mtRi_{1-100}$/$GFP11_{x7}$ were transiently expressed under control of CaMV35S promoter. In addition, the above constructs were co-expressed in reciprocal combinations, i.e. **(c)** FNR_{1-55}/GFP1-10 with $mtRi_{1-100}$/$GFP11_{x7}$ or **(f)** $mtRi_{1-100}$/GFP1-10 with FNR_{1-55}/$GFP11_{x7}$, after *Agrobacterium* co-infiltration into the lower epidermis of *Nicotiana benthamiana*. In all instances, no fluorescence signals were observed. The *squares* highlight areas of the transformed cells as indicated. Scale bars correspond to 20 µm. (For further details, see legend of figure 3.2)

Next, the protein targeting behaviour of several dually targeted proteins was determined by this system. For this purpose, three previously characterized nuclearly encoded organelle proteins were selected from *Arabidopsis thaliana* with proven dual targeting characteristics, namely TyrRS (Tyrosine-tRNA synthetase; At3g02660), GrpE (co-chaperone GrpE1; At5g55200), and PDF (peptide deformylase 1B; At5g14660) (Berglund et al., 2009; Baudisch et al., 2014). The experiments employing eYFP (enhanced Yellow Fluorescent Protein) fusions had shown that all three proteins are targeted to both endosymbiotic organelles, either in comparable amounts ($TyrRS_{1-91}$/eYFP) or preferentially to either mitochondria ($GrpE_{1-100}$/eYFP) or chloroplasts (PDF_{1-100}/eYFP) (Figure 3.4, left panels and section 2.3.3). The respective N-terminal amino acid sequences carrying the organelle transport signals of the candidate proteins were fused to $GFP11_{x7}$ tags and analysed in our system. All three candidates ($TyrRS_{1-91}$/$GFP11_{x7}$, $GrpE_{1-100}$/$GFP11_{x7}$ and PDF_{1-100}/$GFP11_{x7}$) showed targeting to both plastids and mitochondria when co-transformed with the respective organelle-targeted receptors (Figure 3.4). Even plastid targeting of $GrpE_{1-100}$/$GFP11_{x7}$ was clearly visible, which is remarkable considering the vague fluorescence signals obtained in this organelle with the 'standard' fluorescent reporter fusion (Figure 3.4b). Thus, separation of the fluorescence signals for the two organelles into different cells proved to be advantageous to determine the low plastid targeting properties of GrpE.

3.2.2 Multimerization of the GFP11 tag leads to fluorescence signal enhancement in plastids but not in mitochondria

In the initial experiments described above the seven-fold repeat of the GFP11 tag was used. However, the requirement or benefits of such multiple GFP11 tags to enhance fluorescence signals in mitochondria and plastids of plant cells were not systematically assessed. Hence, we have compared the fluorescence signal intensity obtained in the two organelles with seven ($GFP11_{x7}$), three ($GFP11_{x3}$) or a single repeat ($GFP11_{x1}$) of the GFP11 tag when fused to the dually targeted $TyrRS_{1-91}$ peptide as protein transport signal. It turned out that the use of a single GFP11 tag yields only very faint signals in plastids while the $GFP11_{x3}$ and $GFP11_{x7}$ tags significantly enhance the fluorescence signals (Figure 3.5a) supporting the assumption that multiple GFP11 repeats can boost fluorescence signal intensity in this organelle. In contrast, for mitochondria no such correlation of number of GFP11 tag repeats and fluorescence signal intensity was found. Instead, the signal intensities were largely similar for all three constructs

(Figure 3.5b). However, since the fluorescence signals obtained with the single GFP11 tag in mitochondria were brighter than in plastids, they are usually sufficient for proper visualization of organelle.

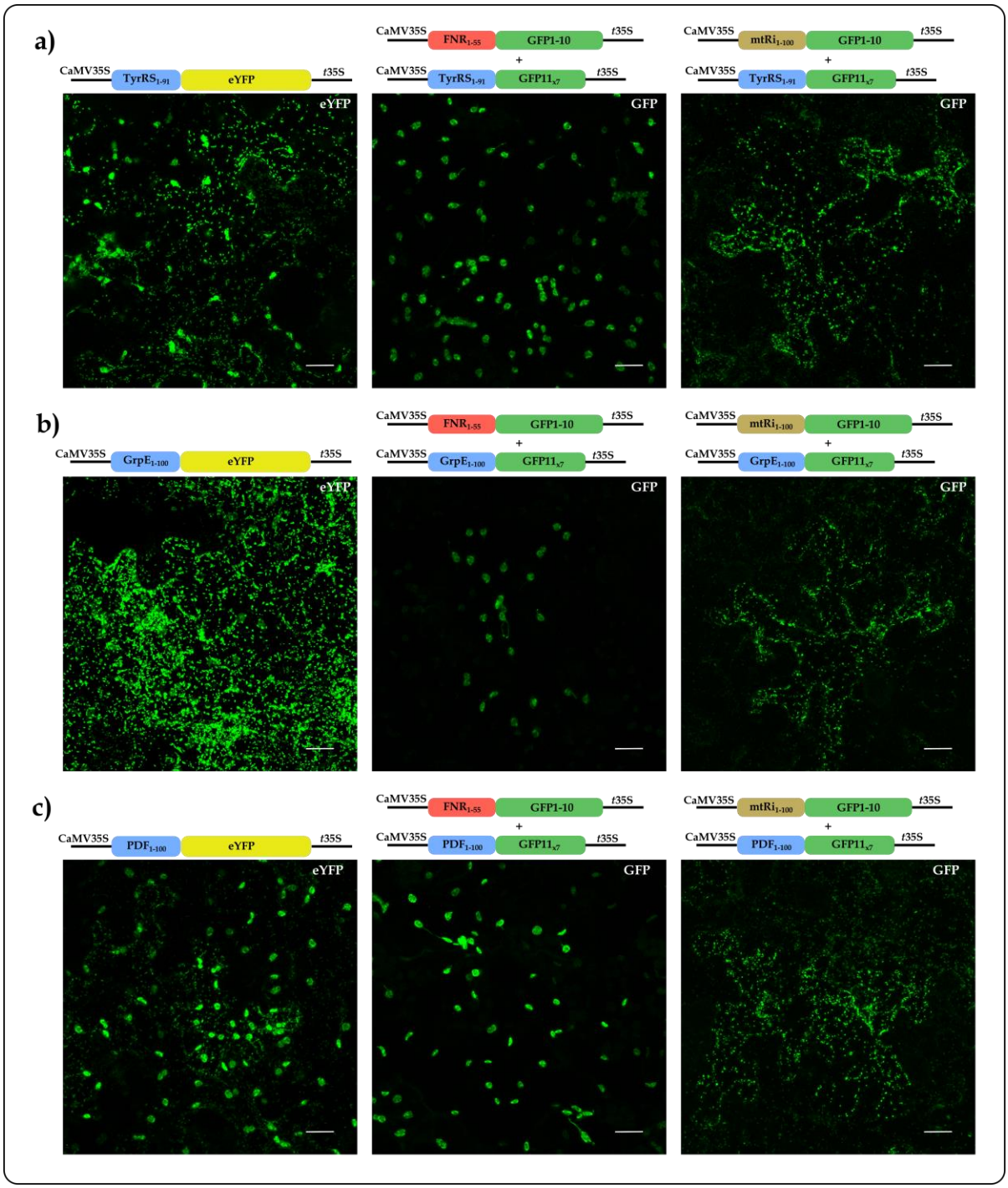


Figure 3.4 Comparison of FP-tagging and *sa*split-GFP approaches to analyze the targeting specificity candidate proteins with proven dual-targeting properties. The coding sequences of **(a)** TyrRS$_{1-91}$/GFP11$_{x7}$, **(b)** GrpE$_{1-100}$/GFP11$_{x7}$ and, **(c)** PDF$_{1-100}$/GFP11$_{x7}$ were transiently expressed with either FNR$_{1-55}$/GFP1-10 (*middle panels*) or mtRi$_{1-100}$/GFP1-10 (*right panels*) via Agrobacterium co-infiltration into the lower epidermis of *Nicotiana benthamiana* leaves and analyzed by CLSM. For comparison, the respective eYFP fusions namely **(a)** TyrRS$_{1-91}$/eYFP, **(b)** GrpE$_{1-100}$/eYFP, and **(c)** PDF$_{1-100}$/eYFP were transiently expressed as well and analyzed by CLSM (*left panels*). For further details, see the legend of Figure 2. Scale bars correspond to 20 µm.

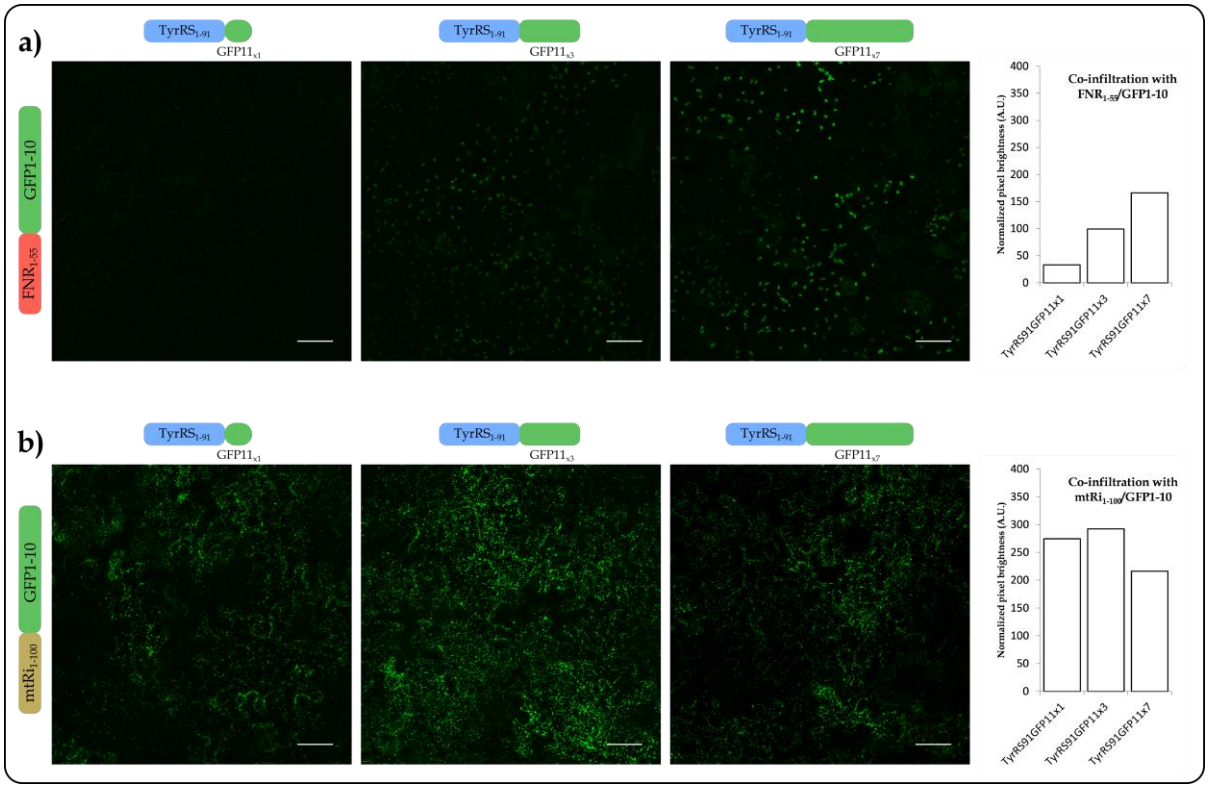


Figure 3.5- Effect of multiple GFP11 tags on fluorescence signal intensity. The gene coding sequence of the dual targeting transit peptide $TyrRS_{1-91}$ was fused with different GFP11 variants ($GFP11_{x1}$, $GFP11_{x3}$ and $GFP11_{x7}$) and co-expressed with either **(a)** FNR_{1-55}/GFP1-10 or **(b)** $mtRi_{1-100}$/GFP1-10 in *Nicotiana benthamiana* lower epidermal cells. Images from GFP channel are presented. The graphs represent normalized average pixel brightness in arbitrary units (A.U.) obtained after subtraction of background fluorescence signals from 9 maximum intensity projected images from two plant replicates (see Figure 3.7 for absolute quantification). Scale bars correspond to 50 µm.

3.2.3 Construction of the PlaMiNGo toolkit

To facilitate high-throughput screening of protein targeting specificity with the *sa*split-GFP technology, we have developed a set of *Golden Gate*-based vectors. Destined to analyze the targeting specificity of candidate proteins to plastids and mitochondria, these vectors facilitate easy cloning of the candidate proteins upstream of single or multiple GFP11 tags. In this PlaMiNGo toolkit, we have used a 'long' 35S promoter (Engler et al., 2014) and *ocs* or *rbcs E9* transcriptional terminators (De Greve et al., 1982; Coruzzi et al., 1984) to control the expression of the chimeric genes (Figure 3.6). Moreover, to avoid the requirement for performing co-transformation using two plasmids, we have utilized a single *T-DNA* expression system (Grefen and Blatt, 2012; Hecker et al., 2015) comprising two gene expression cassettes. This has the advantage that each transformed cell expresses both chimeras simultaneously. Consequently, the final vectors (Figure 3.6-module II) contain expression cassettes for GFP1-10 gene chimeras to be targeted to either plastid or mitochondria by fusion with $mtRi_{1-100}$ or FNR_{1-55} and another for one, three or seven times repeat of GFP11 tags. The latter expression cassette furthermore carries a *ccd*B cassette upstream of the GFP11 tag, which can be replaced by the candidate gene in a

single *Golden Gate* reaction step allowing for background-free selection of positive clones. As a result, six vectors, three destined to analyze potential plastid targeting and three for mitochondria targeting analysis, were generated (Figure 3.6-module II).

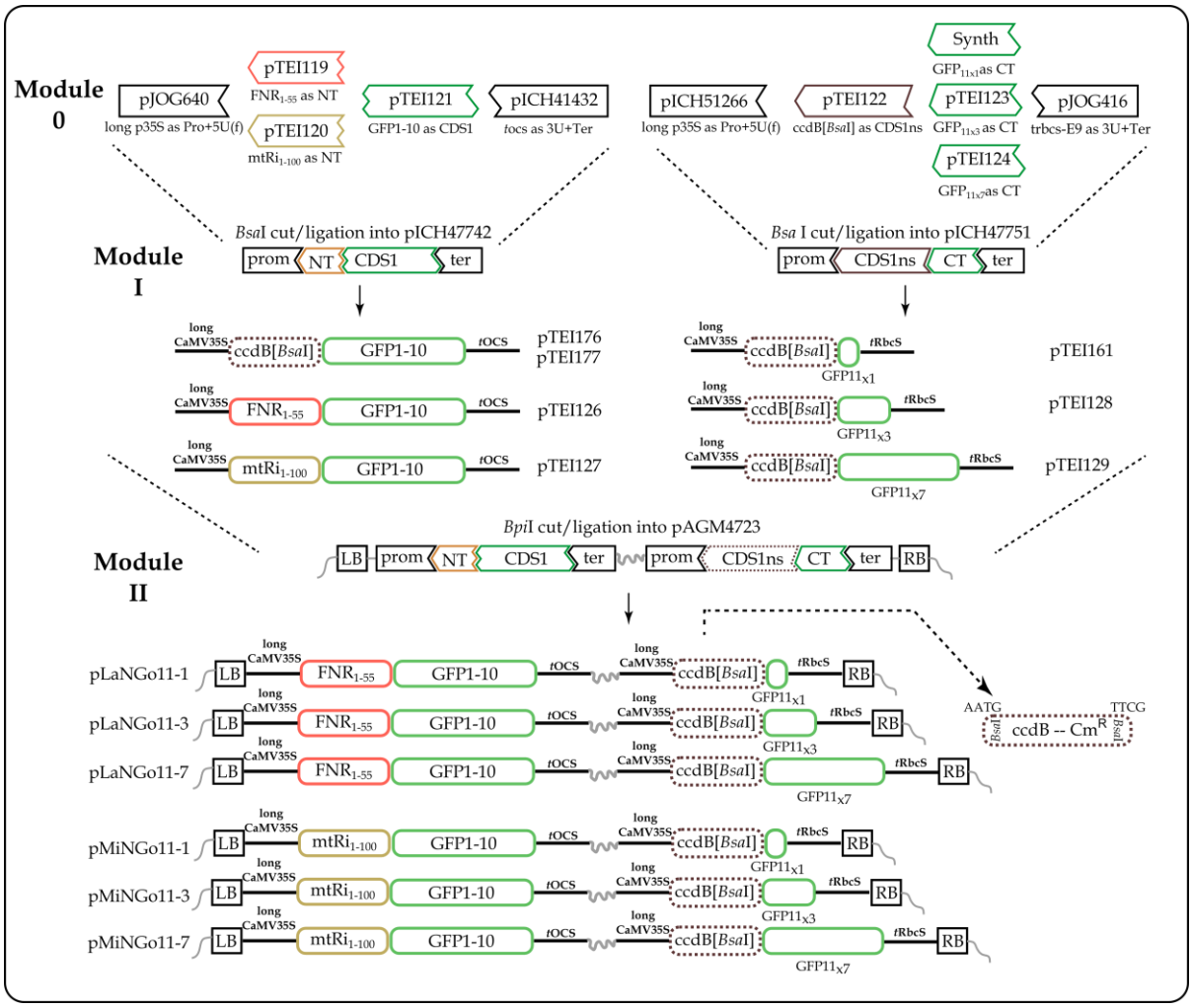


Figure 3.6 Assembly of Golden Gate-based vectors and construction of the PlaMiNGo toolkit. The Modular Cloning system (Weber et al., 2011) and DNA modules from the Plant Parts I and II toolkits (Engler et al., 2014; Gantner et al., 2018) were used for vector assembly. All generated modules and vectors are summarized here. Level 0 modules were generated using gene specific primers (summarized in Table 6.5). Oligonucleotides for $GFP11_{x1}$ fragment were custom synthesized and hybridized before cloning. Five independent vector constructions (pTEI126, pTEI127, pTEI161, pTEI128, pTEI129) were generated in level I cloning steps. Two constructions namely pTEI176 and pTEI177 were generated additionally to extend the toolkit for the analysis of protein targeting to different subcellular compartments. The level I modules were then used to generate final vectors for analysis of protein targeting specificity to plastids (pLaNGo11-1, pLaNGo11-3, pLaNGo11-7) or mitochondria (pMiNGo11-1, pMiNGo11-3, pMiNGo11-7). These final PlaMiNGo vectors carry a *ccdB* negative selection cassette flanked with BsaI restriction site, which should be replaced by candidate gene sequence after *Golden Gate* cloning. (The toolkit was constructed in collaboration with Dr. Johannes Stuttmann, MLU Halle, Germany).

To evaluate the functionality of these vectors, the dual targeting transport signal of TyrRS (N terminal 1-91 amino acids) was cloned upstream of the GFP11 tags of all six vectors and analysed via *Agrobacterium*-mediated transient transformation of *Nicotiana benthamiana* leaves. In all instances, tremendously improved signal intensities were observed for both organelles

compared to the previous co-infiltration experiments with two-vector system (Figure 3.7). Now, even in plastids a single copy of the GFP11 tag was sufficient for reliable detection of the reconstituted GFP fluorescence, although considerable signal enhancement could still be observed with $GFP11_{x3}$ and $GFP11_{x7}$ tags (Figure 3.7a, b). This suggests that the $GFP11_{x7}$ tag, when co-expressed with the plastid targeted GFP1-10 receptor, should allow for detection even of minute amounts of plastid-localized proteins. Likewise, upon imaging of mitochondria the fluorescence signals obtained with the single GFP11 tag were more than 3-fold stronger than the signals obtained in the previous experiments using separate vectors. Still, as observed earlier, no further improvement of signal intensities by multimerization of the GFP11 tag could be obtained in mitochondria (Figure 3.7c, d). Instead, artificial protein aggregates were observed in some cells expressing constructs with multiple GFP11 tags (GFP_{11x3} and GFP_{11x7}) (Figure 3.8).

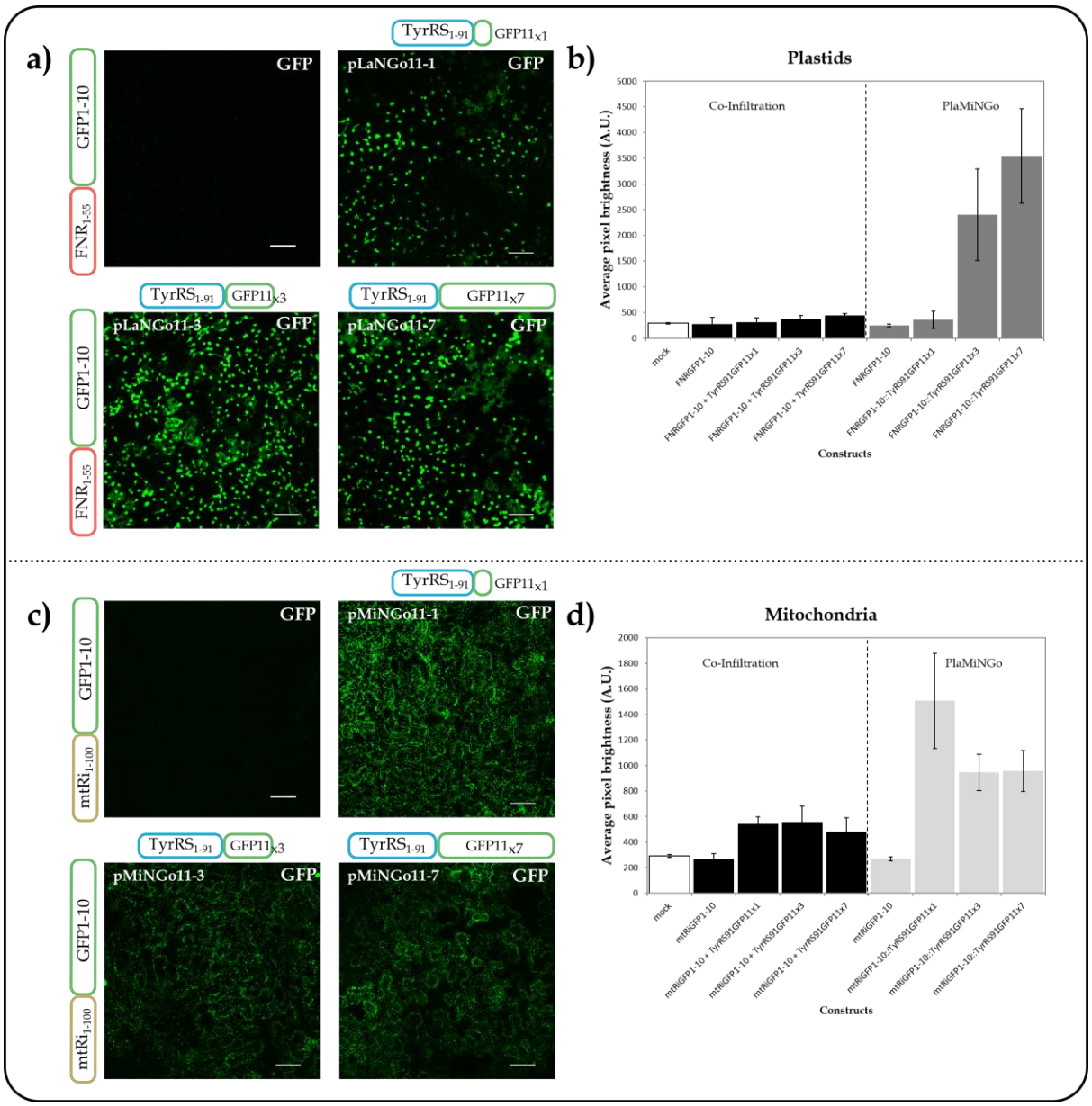


Figure 3.7 Improved fluorescence signals with PlaMiNGo toolkit. The dually targeted $TyrRS_{1-91}$ transit peptide was cloned in the respective PlaMiNGo vectors and used to transform *Nicotiana benthamiana* lower epidermis cells via *Agrobacterium* infiltration. The representative images are showing subcellular localization of $TyrRS_{1-91}$ fused to different version of GFP11 fragments. Images from GFP channel are presented. The charts display a comparison between PlaMiNGo vectors (section 2.2.3) and vectors used for co-infiltration experiments (section 3.2.2). Each bar represents the average pixel brightness in arbitrary units (A.U.) of nine different digital images obtained from 2-3 independent experiments. Image acquisition was done with 20-x objective and 2% of full laser power in several equally separated Z-stacks. Scale bars on microscopy images correspond to 50 µm

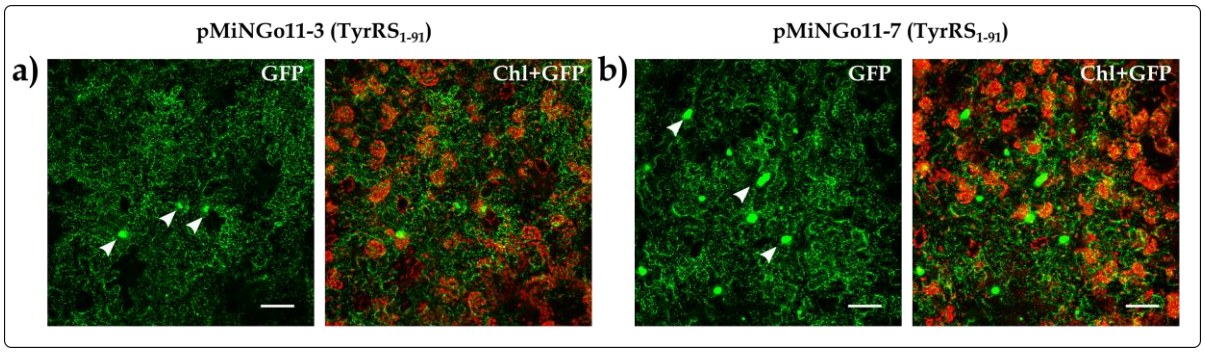


Figure 3.8 Formation of aggresomes by multiple GFP11 repeats. Protein aggregates are often observed (indicated by white arrows) in cells transformed with **(a)** pMiNGo11-3 and **(b)** pMiNGo11-7 vectors carrying three times or seven times GFP repeats at C-terminal of candidate proteins, e.g. $TyrRS_{1-91}$. For quantification purposes, those areas were selected where there were no such aggregates visible. Scale bars correspond to 50 µm

3.2.4 Multicolor imaging of dual protein targeting to two organelles

A further objective of the study was to establish simultaneous multicolor imaging with the *sa*split-FP system. For this purpose, we have modified the GFP1-10 receptor to generate a yellow shifted variant (YFP1-10) using a single amino acid substitution (T203Y) as reported earlier (Kamiyama et al., 2016). To test if this variant can assemble with GFP11 to generate a functional fluorophore inside the organelles, the plastid (FNR_{1-55}) or mitochondria ($mtRi_{1-100}$) transport signals were separately fused to the N-terminus of YFP1-10. These fusions were co-infiltrated with a gene construct encoding the dual targeting transport signal $TyrRS_{1-91}$ fused to either $GFP11_{x1}$ or $GFP11_{x7}$. When imaged with an YFP-specific filter-set, fluorescence signals were solely obtained in mitochondria (with $mtRi_{1-100}$/YFP1-10 and $TyrRS_{1-91}/GFP11_{x1}$) or in plastids (with FNR_{1-55}/YFP1-10 and $TyrRS_{1-91}/GFP11_{x7}$), demonstrating that the YFP1-10 fragment can indeed assemble with GFP11 in both organelles (Figure 3.9a, 3.9b).

Next, we have tested if multicolor imaging, i.e. the simultaneous labelling of plastids and mitochondria with YFP1-10 and GFP1-10, respectively, within the same cell is possible. For this purpose, the gene coding sequence for FNR_{1-55}/YFP1-10 fusion was co-infiltrated with a vector comprising $mtRi_{1-100}$/GFP1-10 and $TyrRS_{1-91}/GFP11_{x1}$. This should result in transformed cells co-expressing the dually targeted GFP11 (via fusion with $TyrRS_{1-91}$) and two different receptors targeted to two different organelles, namely YFP1-10 to plastids (FNR_{1-55}/YFP1-10) and GFP1-

10 to mitochondria (mtRi$_{1-100}$/GFP1-10). Indeed, the resulting transformed cells emitted fluorescence signals of different spectra in the two organelles due to reassembly of the GFP11 tag with both, the GFP1-10 and YFP1-10 receptors within mitochondria and plastids, respectively (Figure 3.9c). However, in all transformed cells a certain degree of "bleed through" of signals, i.e., the appearance of YFP fluorescence signals in GFP channel and *vice versa*, could be detected. The adjustment of the filter-sets to avoid such “bleed through” inevitably led to significant reduction of the fluorescence signal intensity. In summary, the YFP1-10 derivative of *sa*split-GFP system is not yet perfect for multicolor imaging but represents a promising basis for development of such tools.

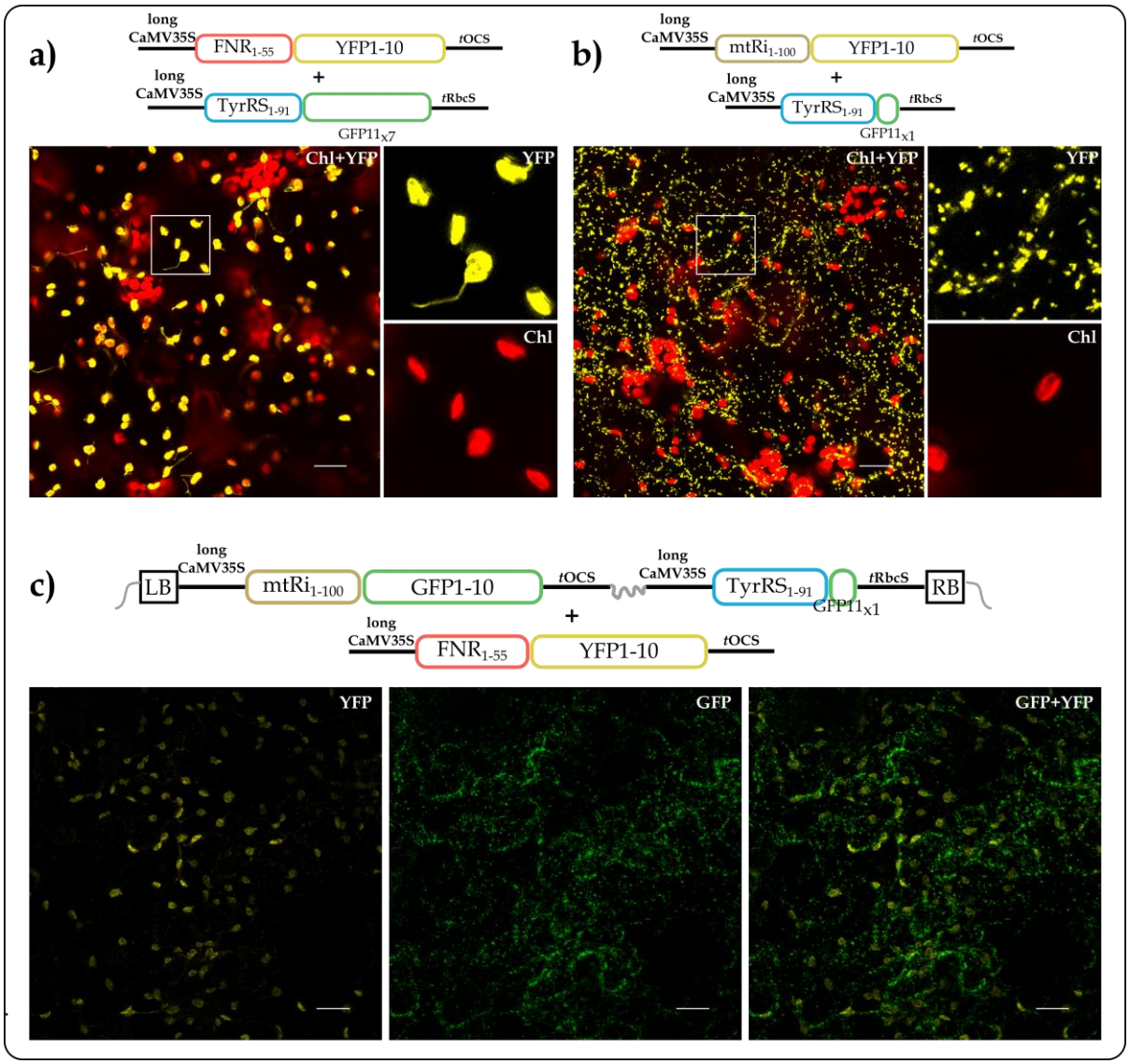


Figure 3.9 Multicolor imaging with *sa*split-YFP. The gene coding sequences of **(a)** FNR$_{1-55}$/YFP1-10 and TyrRS$_{1-91}$/GFP11$_{x7}$, or **(b)** mtRi$_{1-100}$/YFP1-10 and TyrRS$_{1-91}$/GFP11$_{x1}$ were transiently co-expressed in *Nicotiana benthamiana* leaf epidermis and analyzed by CLSM. **(c)** For the purpose of simultaneous multicolor imaging within the same cell, the PlaMiNGO vector comprising mtRi$_{1-100}$/GFP1-10 and TyrRS$_{1-91}$/GFP11$_{x1}$ was co-infiltrated with FNR$_{1-55}$/YFP1-10 resulting in self-assembly of GFP in mitochondria and YFP in plastids. Scale bars correspond to 20 µm.

3.2.5 Analysis of protein targeting specificity with PlaMiNGo

Finally, the suitability of the PlaMiNGo toolkit was examined using eight candidate transit peptides with presumed targeting specificity either to plastids, mitochondria, or to both organelles (Table 1). Three of these proteins, namely Gtred (Monothiol glutaredoxin-S15), GCS (Glycine cleavage system H protein 1) and GAPDH (Glyceraldehyde-3-phosphate dehydrogenase B) had earlier been reported to be dually targeted (Baudisch et al., 2014). Two other candidates, namely FNR and RbcS (small subunit of Rubisco from pea), are well-characterized plastid proteins (Highfield and Ellis 1978; Zhang et al., 2001), while the residual three candidates, namely mtRi, ATPS (ATP synthase subunit beta-3) and CoxIV (Cytochrome *c* oxidase subunit IV) of yeast, are known for their mitochondrial targeting specificity (Maarse et al., 1984; Emmermann et al., 1994; Baudisch et al., 2014). The protein fragments comprising the transport signals of these proteins were cloned as fusions with GFP11 tags into the PlaMiNGo vectors, additionally comprising the organelle targeted receptor fusions FNR_{1-55}/GFP1-10 or $mtRi_{1-100}$/GFP1-10.

Five of the eight candidates showed in our assay system the same targeting behaviour as reported in the literature (Table 3.1): mtRi and FNR showed exclusive transport into either mitochondria or plastids, respectively (Figure 3.10a-b) (Rodiger et al., 2011), and the dual targeting candidates GCS, Gtret and GAPDH showed transport into both organelles (Baudisch et al., 2014) (Figure 3.10c-e). However, in the case of GAPDH mitochondrial targeting was rather weak in our assays and could be observed only in few transformed cells. However, it was rather unexpected that the remaining three monospecific candidates, namely $ATPS_{1-100}$, $RbcS_{1-79}$ and $CoxIV_{1-29}$, showed dual targeting in our experiments (Figure 3.11). In the literature, mitochondrial targeting of RbcS has already been described once (Rudhe et al., 2002) but these results were solely based on *in vitro* assays. In addition, dual targeting of the yeast mitochondria presequence, CoxIV, could be assumed considering high degree of freedom of non-plant mitochondria transport signals (Staiger et al., 2009). However, this dual targeting was entirely unexpected for the plant mitochondrial protein, $ATPS_{1-100}$, which never showed any targeting to plastids when analysed with *in vivo* fluorescent protein tagging and *in vitro* protein transport experiments (Baudisch et al., 2014).

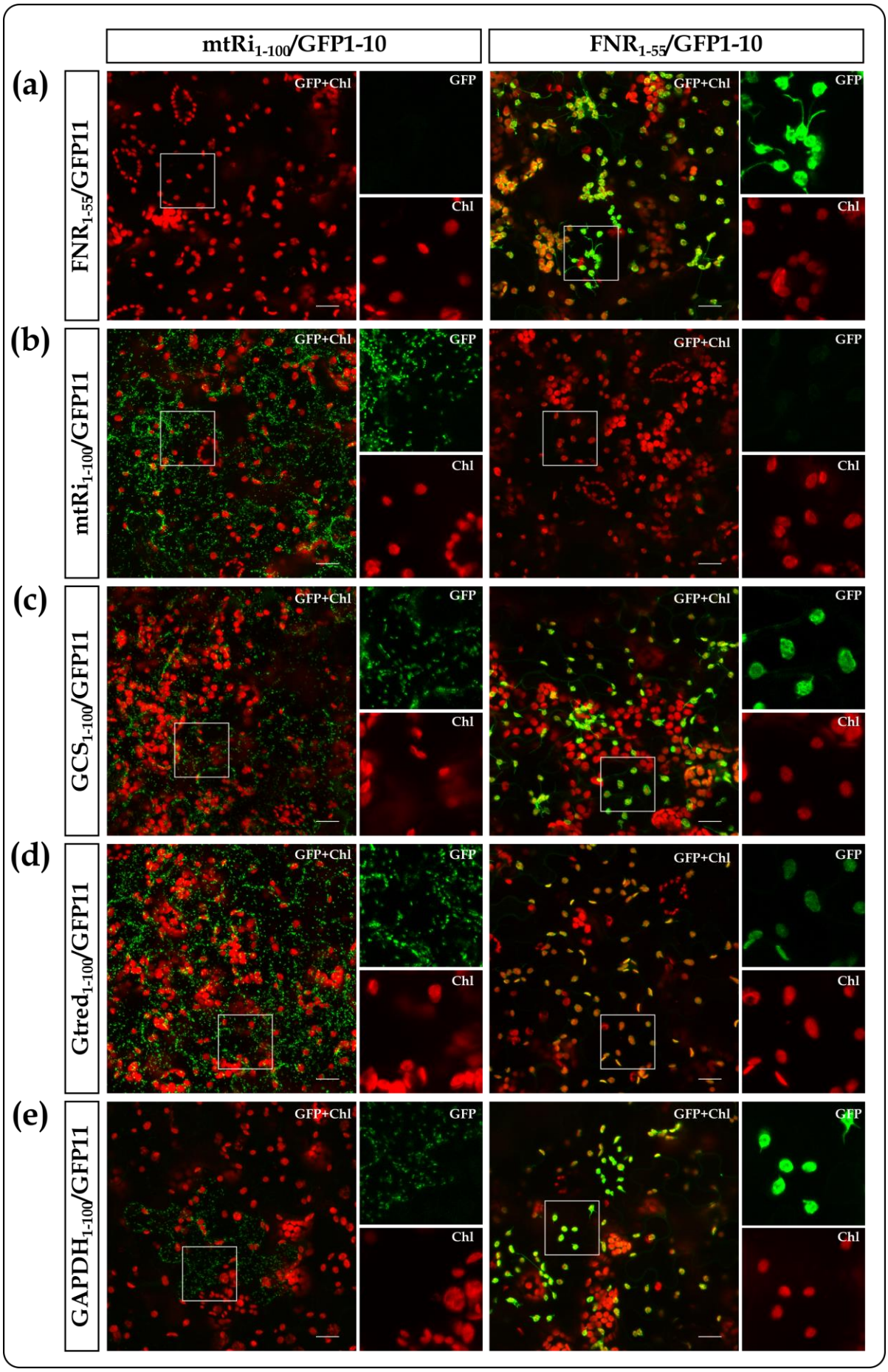

mtRi1-100/GFP1-10
FNR1-55/GFP1-10
(a)
FNR1-55/GFP11
(b)
mtRi1-100/GFP11
(c)
GCS1-100/GFP11
(d)
Gtred1-100/GFP11
(e)
GAPDH1-100/GFP11
GFP+Chl
GFP
Chl

Figure 3.10 Analysis of candidate proteins with the PLaMiNGo toolkit. Gene fragments encoding the transport signals of either **(a)** FNR, **(b)** mtRi, **(c)** GCS, **(d)** Gtred or **(e)** GAPDH, were cloned upstream of either $GFP11_{x1}$ or $GFP11_{x7}$ tags in the two PlaMiNGo vectors comprising either $mtRi_{1-100}$/GFP1-10 (*left panels*) or FNR_{1-55}/GFP1-10 (*right panels*) respectively. The resulting constructs were used to transform the lower epidermis of *Nicotiana benthamiana* leaves via *Agrobacterium* infiltration. Image acquisition of transformed cells was done with 20-x objective in several Z-stacks, which were subsequently stacked for maximum intensity projection. Representative cells are presented as overlay images of the chlorophyll channel (displayed in red) and the GFP channel (displayed in green). The strong chlorophyll signals in the background are derived from the larger chloroplasts of untransformed mesophyll cells underneath the epidermal cell layers. The *squares* highlight areas of the transformed cells shown in higher magnification separately for the chlorophyll channel and the GFP channel as indicated. Scale bars correspond to 20 µm.

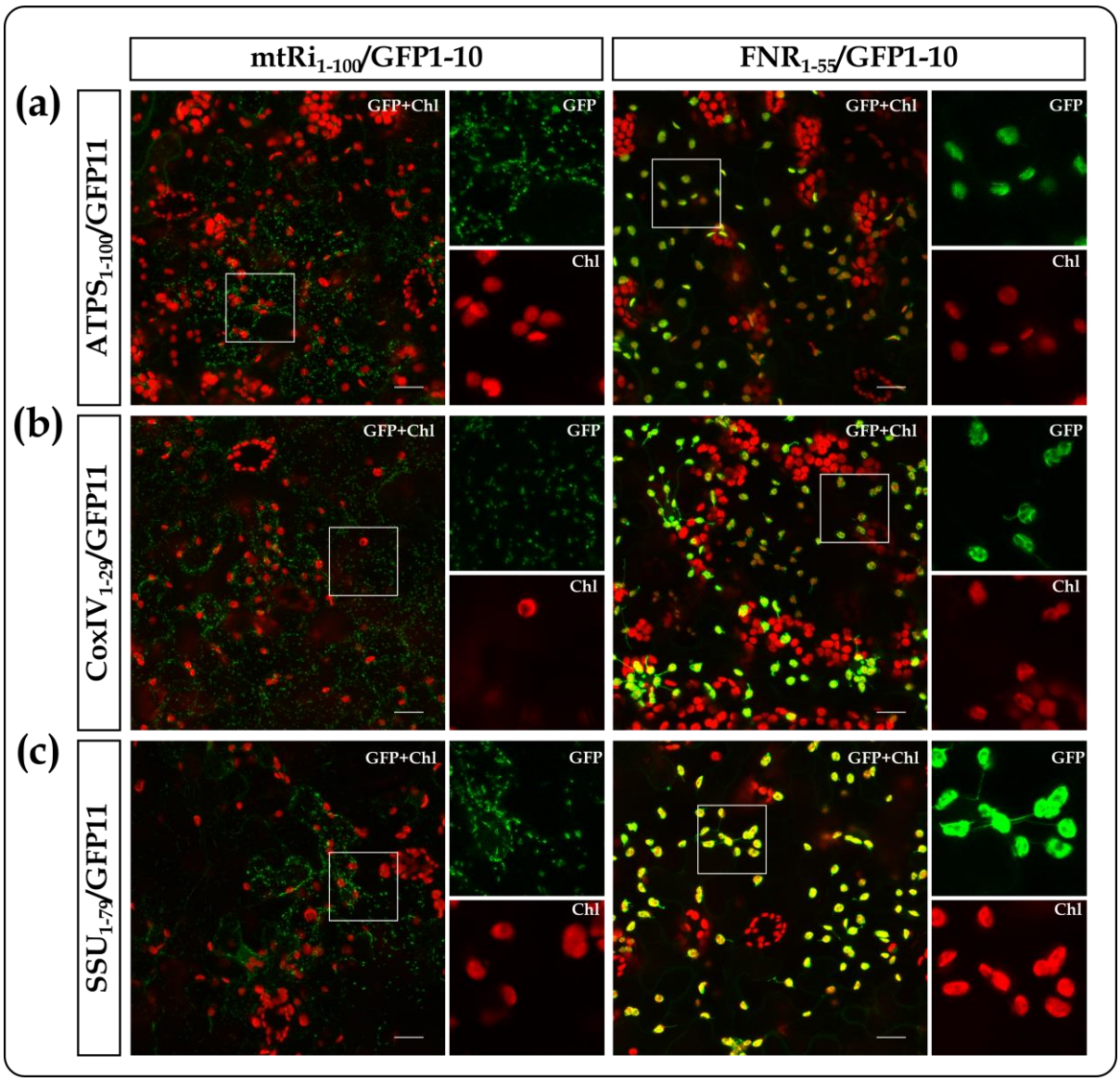


Figure 3.11 Dual localization of presumed monospecific candidate proteins. Gene fragments encoding the transport signals of either **(a)** ATPS, **(b)** CoxIV or **(c)** SSU, were cloned upstream of the $GFP11_{x7}$ tag in the two PlaMiNGo vectors comprising either $mtRi_{1-100}$/GFP1-10 (*left panels*) or FNR_{1-55}/GFP1-10 (*right panels*). The resulting constructs were used to transform the lower epidermis of *Nicotiana benthamiana* leaves via *Agrobacterium* infiltration. For further details, see the legend of Figure 3.10. Scale bars correspond to 20 µm.

Table 3.1 Subcellular localization of candidate proteins as determined with different approaches.

Candidate	*sa*split-GFP	*in vitro* import¶	FP Localization	Gene Accession
FNR	Chloro	Chloro[2]	Chloro[2]	M86349.1
mtRi	Mito	Mito[2]	Mito[2]	X79332.1
GCS	Mito and Chloro	Mito and Chloro[2]	Mito and Chloro[2]	At2g35370
GAPDH	Mito and Chloro	Mito and Chloro[2]	Chloro[2]	At1g42970
Gtred	Mito and Chloro	Mito and Chloro[2]	Mito and Chloro[2], Chloro[4], Mito[5]	At3g15660
CoxIV	Mito and Chloro	NA	Mito[6]	SGD:S000003155
ATPS	Mito and Chloro	Mito[2]	Mito[2]	At5g08680
SSU	Mito and Chloro	Mito and Chloro[1]	Chloro[3]	XM_016585367

¶Import studies were performed with the authentic precursor proteins; *NA*- not avialable ; [1]Rudhe et al., 2002; [2] Baudisch et al., 2014; [3]Nelson et al., 2007; [4]Cheng, 2008; [5]Moseler et al., 2015; [6]Köhler et al., 2003; *Mito*- mitochondria; *Chloro*- Chloroplasts or plastids.

3.2.6 Phenotype complementation confirms the plastid targeting properties of nuclearly encoded proteins

These unexpected results demanded for independent confirmation. Thus, to re-evaluate the plastid targeting properties of ATPS a phenotype complementation approach using *immutans* (*im*) mutants of *Arabidopsis thaliana* was established in this study. These *im* mutants are devoid of a plastid targeted protein namely PTOX (plastid terminal oxidase) and show variegated (green and white-sectored) leaf phenotype when grown under normal daylight (Carol et al., 1999; Aluru et al., 2001). The plastid targeting of PTOX is mediated by a 56 AA long, N-terminal cleavable transit peptide. In order to adapt this system to determine plastid targeting specificity of candidate transit peptides, the authentic transit peptide of PTOX was replaced with the transport signals of our candidate proteins. Hence the complementation of *immutans* phenotype should, in principle, confirm the plastid targeting specificity of a given candidate transit peptides (Figure 3.12).

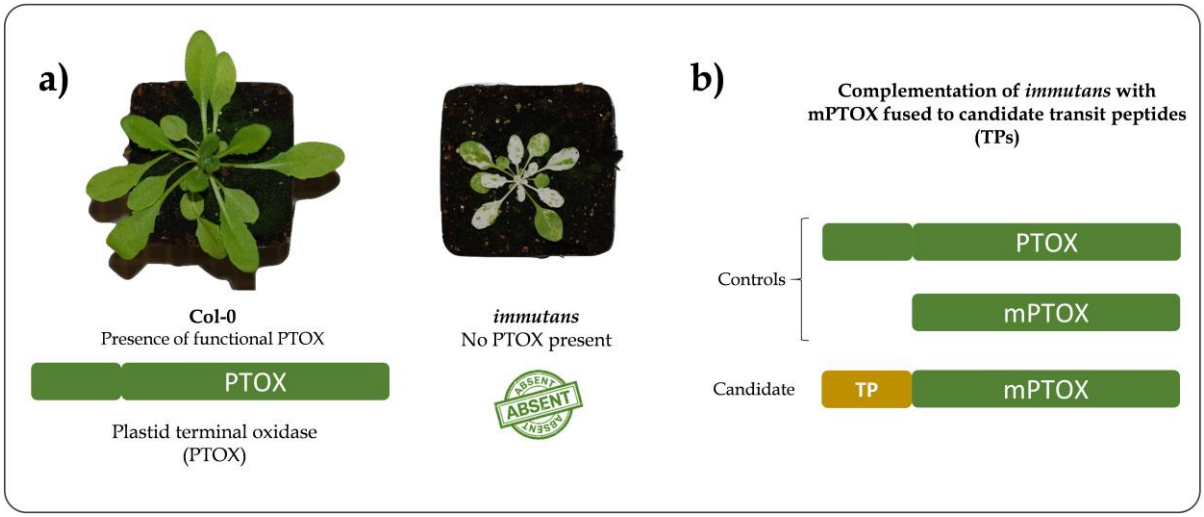


Figure 3.12 Principle of phenotype complementation approach using *immutans*. **(a)** The *immutans* mutants of *Arabidopsis thaliana* lack plastid targeted PTOX protein and thus show variegated (white sectored) leaf phenotype. **(b)** The variegated phenotype can be complemented only if the expressed PTOX protein is transported into chloroplasts. Thus, the transit peptide of PTOX was replaced with the candidate transit peptide and used to transform *immutans* plants. A phenotype complementation should confirm plastid targeting properties of candidate transit peptide.

To establish that the plastid targeting of PTOX is indeed necessary for complementation of *immutans* phenotype, constructions encoding either the authentic PTOX precursor or the transit peptide-free mature PTOX protein (mPTOX) were generated. These gene chimeras were expressed under the control of the CaMV35S promoter in the *immutans* mutant plants. As expected, the authentic PTOX precursor was able to complement the variegated phenotype while the expression of mPTOX could not (Figure 3.13c-d). This provided a proof of principle that the *immutans* mutant can be utilized as a tool to assess plastid targeting specificity of candidate proteins.

Further, to analyze plastid targeting properties of ATPS, the gene sequence coding for authentic transit peptide of PTOX was replaced with the gene sequence coding for amino-terminal 100 AA residues of ATPS generating *ATPS$_{1-100}$/mPTOX* gene chimera. Additionally, three previously characterized dually targeted proteins namely GCS, GrpE and Ef-Tu were included in this analysis. The gene sequence coding for amino-terminal 100 AA residues of these proteins were fused to gene sequence coding for mPTOX generating *GCS$_{1-100}$/mPTOX*, *GrpE$_{1-100}$/mPTOX* and *Ef-Tu$_{1-100}$/mPTOX* gene chimeras. To recap, in stable transgenic lines GCS$_{1-100}$/eYFP accumulated to both plastid and mitochondria while the plastid targeting of GrpE$_{1-100}$/eYFP and Ef-Tu$_{1-100}$/eYFP was not observed (see Section 2.2.3). The latter was also the case with ATPS$_{1-100}$/eYFP, which show exclusive mitochondria targeting in stable transgenic lines with some additional protein aggregates (Figure 3.14a).

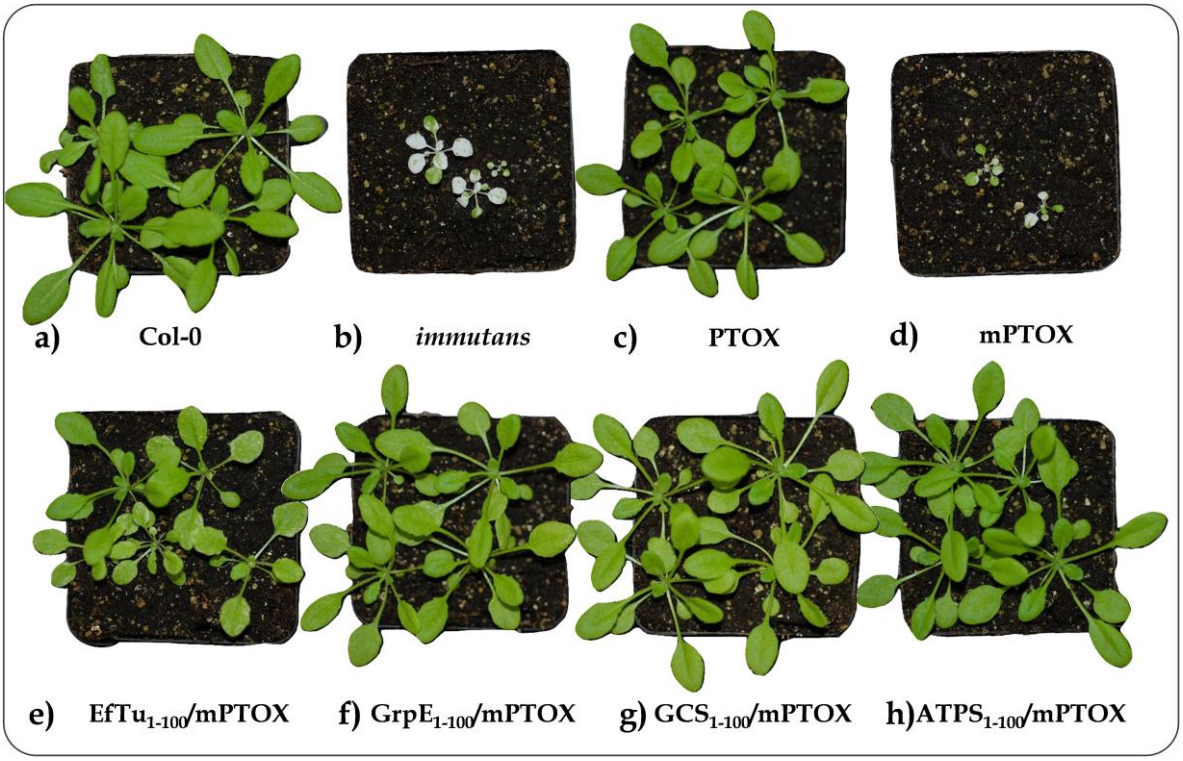


Figure 3.13 Complementation analysis of the *immutans* mutant of *Arabidopsis thaliana*. Phenotype of **(a)** wild type Col-0 and **(b)** *immutans* mutant plants. Gene constructions encoding either **(c)** the precursor PTOX (PTOX), **(d)** the mature PTOX (mPTOX) or PTOX fused to the transit peptides of the candidate proteins **(e)** EfTu, **(f)** GrpE, **(g)** GCS or **(h)** ATPS, were used to transform *immutans* mutant plants. At least three independently transgenic lines were examined for variegated phenotype complementation of the variegated *immutans* phenotype. The photographs show representative 5 weeks old T2 generation plants.

As expected, the GCS_{1-100}/mPTOX was able to complement the variegated phenotype. Remarkably, also the expression of $ATPS_{1-100}$/mPTOX, $GrpE_{1-100}$/mPTOX and Ef-Tu_{1-100}/mPTOX resulted in phenotype complementation (Figure 3.13e-h). However, in case $ATPS_{1-100}$/mPTOX and Ef-Tu_{1-100}/mPTOX not all analysed transgenic lines show complete phenotype complementation suggesting that plastid targeting of these proteins is, in principle, possible but apparently less efficient than with typical plastid targeted transit peptides. These results clearly underline the basic plastid targeting properties of mitochondrial protein ATPS and thus re-confirm the results obtained with the *sa*split-GFP technology established in this study. The expression of chimeric mPTOX in transgenic plants was confirmed via immunoblot assays with the PTOX specific antibodies. Variable sizes of mature PTOX were detected in different transgenic lines indicating different processing sites of transit peptides. Surprisingly, in case of transgenic plants expressing $ATPS_{1-100}$/mPTOX and Ef-Tu_{1-100}/mPTOX, no detectable PTOX specific signals were visible suggesting a potential degradation of mPTOX in this case (Figure 3.14b)

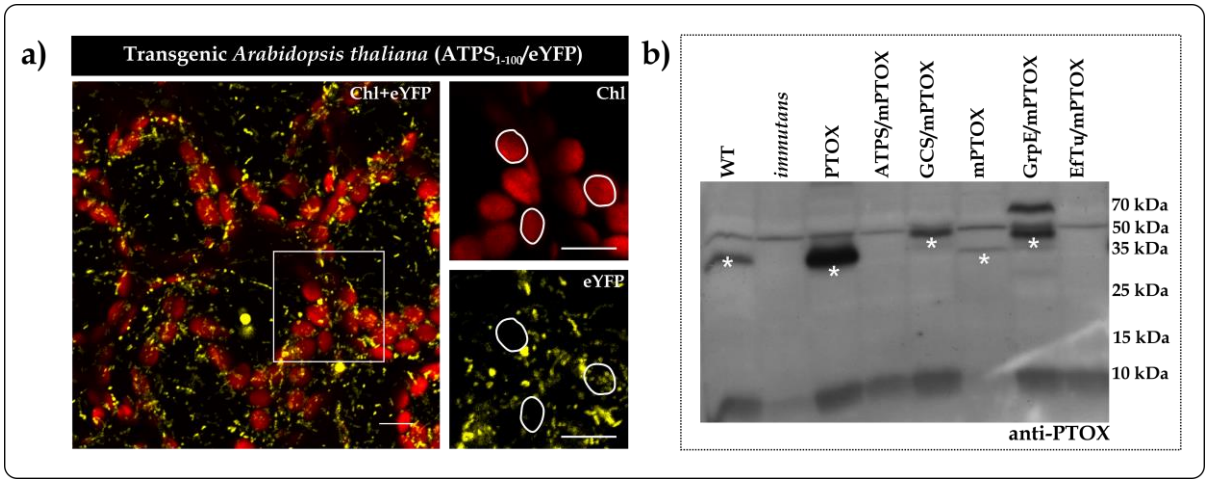


Figure 3.14 (a) Subcellular localization of ATPS. Subcellular localization of $ATPS_{1-100}$/eYFP was determined by confocal laser scanning microscopy in leaf tissue of transgenic *Arabidopsis thaliana* lines. Overlay picture of both the eYFP channel (displayed in yellow) and the chlorophyll channel (displayed in red) are shown. Separate images of the two channels are displayed at higher magnification as indicated. A few chloroplasts are encircled with a while line for better display of signals. The scale bars correspond to 10 µm. (b) Western blot analysis of transgenic *immutans* lines. Total proteins were isolated from 1 mg fresh weight of transgenic leaf tissue, and equal amounts were electrophoresed through 15% SDS-polyacrylamide gels and then transferred to PVDF membranes. The membranes were treated with an antibody to PTOX and visualized by the ECL system. White asterisks indicate the PTOX specific signals.

3.3 Discussion

The goal of this study was to assess the suitability of the signal enhanced *sa*split-GFP system to determine the targeting specificity of nuclearly encoded organelle proteins and to develop tools for rapid cloning and subsequent analysis of their targeting behavior. The application of the *sa*split-GFP system, as demonstrated in this study, provides a novel toolbox to determine the targeting properties of candidate proteins with high sensitivity.

3.3.1 Selective imaging and fluorescence signal enhancement with sasplit-GFP technology

(a) Selective Imaging

The selective imaging of organelles is one of the major advantages of the *sa*split-GFP system in comparison to 'standard' fluorescent protein tagging approaches. The requirement for the presence of the non-fluorescing GFP1-10 receptor in a specific subcellular location is the key for selective imaging (Kaddoum et al., 2010). In this study, two transport signals, namely FNR_{1-55} and $mtRi_{1-100}$, were selected for localization of the receptor specifically within two subcellular locations, the plastid stroma and the mitochondrial matrix, respectively. As a result, fluorescence signals will appear only if the GFP11 tagged protein is completely imported into the same subcellular location and not if a protein is merely binding to the organelle surface. This was otherwise difficult to distinguish with FP-based approaches, specifically for mitochondria due to their small size.

(b) Fluorescence signal enhancement

The self-assembling split-GFP molecules have been reported to produce fluorescent signals of lower intensity than 'standard' fluorescent proteins (Kökker et al., 2018). This problem can be circumvented with the use of multiple GFP11 tags. However, the two organelles respond differently to this modification. On one hand, the signal enhancement with multiple GFP11 tags works well in plastids, on the other hand, fluorescence signal enhancement could not be observed in mitochondria (Figure 3.5). One possible reason might be the size difference between these organelles. Plastid volume is comparatively higher than mitochondria and proteins are more dispersed in plastid stroma. Consequently, the chances for self-assembly of *sa*split-GFP fragments in this organelle are lower than in mitochondria and thus fluorescence signal enhancement could be observed by increasing number of GFP11 tags in plastids. Furthermore, differences in the physicochemical properties of the two organelles, e.g. pH, might likewise contribute to this observation.

The use of more efficient gene regulatory elements, i.e. promoter and terminator, in the PlaMiNGo toolkit also led to significant enhancement in fluorescence signal intensity. However, in combination with multiple GFP11 tags such increased gene expression can lead to the formation of aggregations in transformed cells, particularly if these tags are combined with mitochondria targeting transport signals. The intrinsic property of the *sa*split-GFP fragments to form dimers and aggregates (Cabantous et al., 2005) and comparatively less efficient unfoldase activity of the mitochondrial protein translocation machinery (Agarraberes and Dice 2001) could be one of the possible reasons for this phenomenon. Since, the protein unfolding prior to translocation is apparently more efficient in plastids, the multiple GFP11 tags can efficiently be imported into this organelle. On the other hand, high expression of plastid targeted GFP1-10 alone could result in appearance of faint fluorescence signals, probably due to formation of dimers. In most instances, these faint signals are clearly distinguishable from 'actual' fluorescence signals obtained via self-assembly of *sa*split-GFP fragments but still the organelle targeted GFP1-10 alone should be considered as an experimental control to avoid the misinterpretation of results.

3.3.2 High sensitivity of *sa*split-fluorescence protein system

Fluorescent signal enhancement in combination with selective imaging makes the *sa*split-GFP system highly sensitive with respect to targeting specificity determination. In consequence, dual targeting of several proteins was newly detected with the PlaMiNGo toolkit developed here, which had previously been missed due to inherent limitations of 'standard' fluorescent protein tagging approaches. For example, GAPDH shows dual targeting with the *sa*split-GFP system, in line with the results of *in vitro* import experiments (Baudisch et al., 2014). In contrast, with 'classical' *in vivo* approaches using FP-tagging, GAPDH appeared to be solely transported into plastids (Baudisch et al., 2014). Similarly, as shown here, the transit peptide of RbcS is able to translocate the GFP11 tag into mitochondria, but this property remained undetected with the FP-tagging approach. Remarkably, such mitochondria targeting properties of the RbcS transit peptide were also found in a recent study employing sulfadiazine-resistant plants (Tabatabaei et al., 2018). In contrast, dual targeting of the transport signal of yeast CoxIV had never been reported earlier. Such dual targeting of yeast mitochondrial transport signal might be a consequence of the fact that yeast does not contain plastids and thus the transport signal of yeast mitochondria had not 'learned' to distinguish between the two endosymbiotic organelles (Staiger et al., 2009). The plastid targeting of a yeast mitochondria transport signal has also been reported earlier (Huang et al., 1990). However, such dual targeting was most unexpected for the transport signal of ATPS because neither *in vitro* nor *in vivo* approaches gave any hint of the plastid targeting properties of

this transport signal (Baudisch et al., 2014). Even transgenic plants expressing $ATPS_{1-100}$/eYFP did not show any plastid localization (Figure 3.14a). The fact, that it was clearly detectable here and that this result could be independently confirmed by a phenotype complementation approach using *immutans* mutants (Figure 3.13h), underlines the sensitivity of the *sa*split-GFP technology.

3.3.3 Modularity of PlaMiNGO toolkit

The vector toolkit is based on the principle of modular cloning (Weber et al., 2011). Hence, the components of the PlaMiNGo toolkit can easily be rearranged for *in vivo* imaging of proteins targeted to various other subcellular compartments. The vectors constructed in module I (Figure 3.6) are binary vectors and can be utilized for plant cell transformation via *Agrobacterium* or via several other methods, e.g. protoplast transformation or particle bombardment. The gene of interest can be cloned upstream to one of the GFP11 tags with a single *Golden Gate* cloning reaction. Similarly, GFP1-10 can be targeted to the different subcellular or even sub-organellar compartments via cloning of the specific transport signal N-terminally in a *Golden Gate* 'ready' vector (pTEI176 or pTEI177). These vectors carry a ccdB negative selection cassette upstream of GFP1-10 with *Bsa*I restriction sites A|ATG at the 5´ end and T|TCG at the 3´ end. Consequently, the two vectors carrying GFP11 and GFP1-10 gene chimeras should be co-expressed in a single cell in order to determine protein targeting specificity to the organelle of interest.

3.3.4 Phenotype complementation as an alternative approach?

The approach of phenotype complementation using *immutans* mutants proved to be more sensitive than 'standard' FP-tagging approaches to determine plastid targeting specificity of some of the transit peptides, e.g. ATPS. The reason behind this high sensitivity is that only 3% of total PTOX level is sufficient to retain the wild type phenotype (Fu et al., 2009). Thus, even if only a few PTOX molecules are transported into plastids, the phenotype complementation should be achieved. Besides this, it is known that PTOX is required during early stages of chloroplast biogenesis (Foudree et al., 2012). In this case, even if a chimeric protein is transported into plastids during seed germination and degraded at the later stages of plant development, the phenotype should still be recovered. The latter might be the case with Ef-Tu and ATPS, where no PTOX protein was detected in Western blots while the transgenic *immutans* lines displayed a phenotype similar to the wild type plants. In conclusion, this method seems appropriate to examine plastid targeting specificity of nuclearly encoded proteins while not suitable to determine the accumulation of proteins in this organelle.

A similar approach can be utilized to study mitochondrial targeting specificity of proteins. The *starik* mutant of *Arabidopsis thaliana* could be a suitable system for this purpose (Kushnir et al., 2001). The stunted growth and chlorotic leaves phenotype of *starik* mutants, due to the absence of

a mitochondrial targeted ABC transporter Sta1, can aid phenotyping-based assessment of protein targeting specificity *in vivo*. While these phenotype complementation approaches could provide an alternative for targeting specificity determination, one has to be aware of the consequences of this approach. For example, slow growth-rate of mutants and appearance of intermediate phenotypes might significantly delay the analysis. Besides this, the potential counter-selection of stably transformed mutant plants might lead to artefactual phenotype complementation. Nevertheless, this method is suitable for studying protein transport into organelles and should be used as a complementing approach, not as an ‘alternative’ appraoch.

Chapter 4- Determining targeting specificity of chloroplast twin arginine translocase subunit A (TatA)

4.1 Background

The thylakoid membranes of Chloroplast retains an evolutionarily conserved protein transport system that facilitates translocation of fully folded proteins into the thylakoid lumen, termed as the twin-arginine translocase (Tat) (reviewed in Müller and Klösgen 2005; New et al., 2018). The Tat mediated protein translocation is energized via a proton motive force (PMF) and does not require any additional energy source (Cline et al., 1992). In higher plants, three nuclearly encoded plastid proteins namely TatA, TatB, and TatC constitute a functional Tat machinery and all three components are essential for efficient protein translocation (Cline and Mori 2001; Motohashi et al., 2001; Hauer et al., 2013; Wang et al., 2016). Two of these protein subunits, TatB and TatC, interact with each other to form an oligomeric TatBC-receptor complex that recognizes the N-terminal signal sequence of substrate proteins and facilitates their transport into thylakoid lumen with the help of the third protein subunit, TatA. The role of TatA in this functional Tat machinery is not yet fully understood, however, one of the most popular hypotheses suggests that TatA transiently weakens the lipid bilayer surrounding the fully assembled TatBC-substrate protein complex and facilitates the translocation of a substrate protein (Hou et al., 2018). Presence of two arginine residues within the signal sequence is a unique feature of Tat substrate proteins and explains the nomenclature of twin-arginine translocase (Tat) (Chaddock et al., 1995).

Many prokaryotes, including cyanobacteria and α-proteobacteria, also possess a similar Tat machinery in their plasma membranes (Barnett et al., 2011; Nuñez et al., 2012). Consequently, it could possibly be assumed that the thylakoid Tat machinery has evolved from the cyanobacterial counterpart. Interestingly, no such functional Tat machinery has been characterized so far in the mitochondria of higher plants. The presence of mitochondrially encoded TatC like (mtTatC) transcripts was reported a while ago (Sünkel et al., 1994; Braun and Schmitz 1999). However, the lack of satisfactory evidence supporting the presence of a TatC like protein in mitochondria limited further investigation. More recently, Carrie and colleagues (2016) have demonstrated the presence of mtTatC protein and a newly identified TatB-like protein (At5g43680) in

mitochondria of *Arabidopsis thaliana*. This first-hand observation proposed a possibility of a 'minimal' Tat machinery in plant mitochondria consisting of TatB and TatC, where TatB performs a bifunctional role (Carrie et al., 2016). Contrary to this, the so-called 'minimal' Tat systems known hitherto, e.g. in gram-positive bacteria and in jakobid *Andalucia godoyi*, are composed of TatA and TatC, where TatA performs bifunctional role (Jongbloed et al., 2004; 2006; Petru et al., 2018).

In mitochondria, certain inner envelope proteins, for example Rieske iron-sulphur protein (mtRi or Rip, a subunit of cytochrome bc1 complex), need to be translocated from matrix to inner membrane in a fully folded conformation. Thus, a hunt for specific translocation machinery that is capable to translocate folded proteins across the inner mitochondrial membrane was ongoing until Wagener et al. (2011) demonstrated the presence of an exclusive protein translocation system that interact with mtRi in yeast mitochondria. This specific translocase consists of AAA-ATPase Bcs1 protein homo-hexamers forming a BCS1 (cytochrome bc1 synthase 1) complex at inner membrane of mitochondria to facilitate translocation and topogenesis of folded mtRi into cytochrome bc1 complex. Contrastingly, the functional Bcs1 domain of this protein is absent in higher plants and *Arabidopsis thaliana* AAA-ATPase protein do not interact with corresponding mtRi (Zhang et al., 2014). Hence, the pathway that facilitates translocation and topogenesis of folded mtRi in plant mitochondria is still enigmatic.

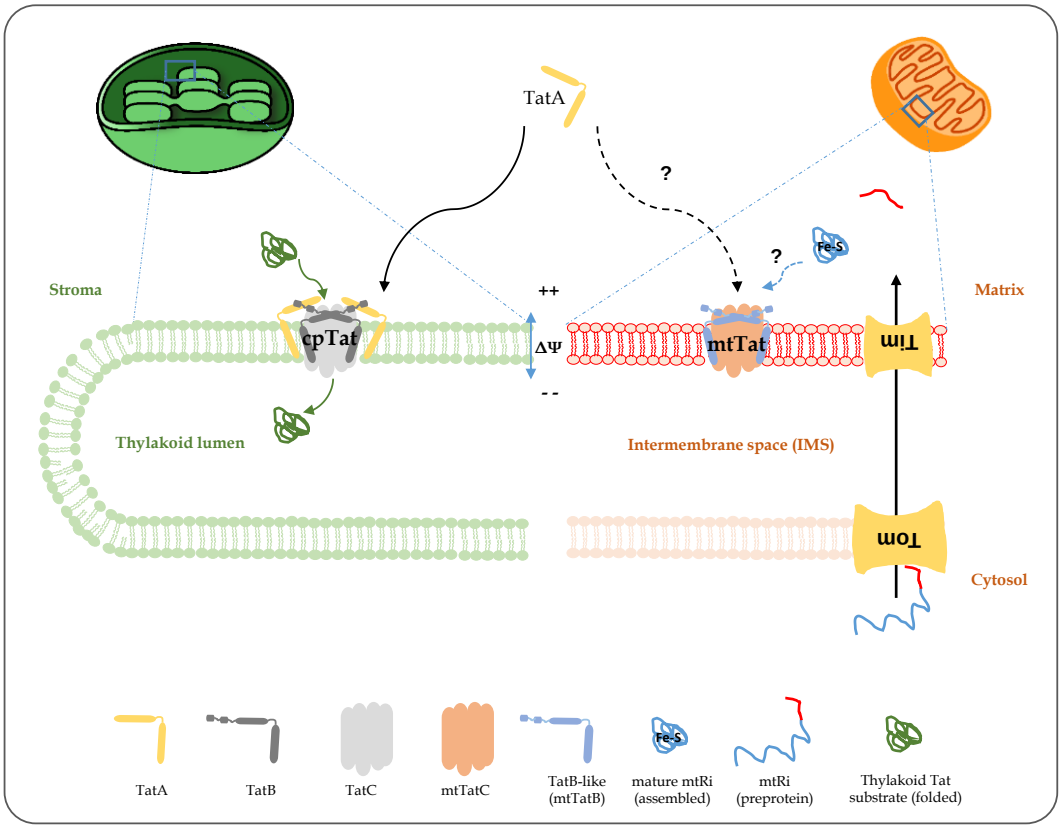


Figure 4.1- Schematic representation of thylakoid Tat machinery and hypothetical model of potential Tat machinery in mitochondria. The thylakoid Tat machinery **(left)** is composed of three chloroplast targeted protein subunits namely TatA, TatB and TatC and facilitate transport of fully folded proteins across and into the membrane against a proton motive force. Similarly the proposed Tat machinery of mitochondria **(right)** is presumably composed of a mitochondrial encoded mtTatC and mitochondria targeted TatB-like protein (Carrie et al., 2016) while the function of chloroplast TatA in the translocase assembly is hypothesized in this study. In this case, mitochondrial Rieske iron-suplhur protein (mtRi) could serve as potential substrate. The targeting of mtRi into mitochondria matrix takes place via general import pathway while its topogenesis in cytochrome *bc*1 complex at inner mitochondrial membrane could possibly be assisted by a mitochondrial Tat translocase.

In chloroplasts, the translocation and topogenesis of an mtRi homolog protein, namely chloroplast Rieske iron-sulphur protein (cpRi), into cytochrome b6f complex is facilitated by a three-component thylakoid Tat machinery. Thus, it could well be that a similar Tat machinery exist in plant mitochondria that facilitate translocation of mtRi. The existence of TatB-like protein and mtTatC in plant mitochondria fits in this context. However, no mitochondria targeted or mitochondrially encoded paralog of TatA has been identified so far. One possibility could be that the nuclearly encoded thylakoid TatA possesses dual-targeting properties and is targeted to both plastids and mitochondria (Figure 4.1). Thus, in order to examine the targeting specificity of TatA protein from two model plants, *Arabidopsis thaliana* and *Pisum sativum*, subcellular localization studies were performed via fluorescent protein tagging and microscopy based approaches. Additionally, *knock-out* mutant of *Arabidopsis thaliana TatA* was generated using Crispr/Cas9 based gene deletion approach (Ordon et al., 2016) to examine the effect of *TatA* deletion on plant phenotype.

4.2 Results

4.2.2 Subcellular localization of *Arabidopsis thaliana* TatA

To determine subcellular protein localization, the gene constructs carrying coding sequence for either N-terminal 100 amino acids (AA) or full length precursor TatA fused to eYFP (*At*Tat$_{1-100}$/eYFP or *At*TatA/eYFP) were expressed under control of CaMV35S promoter in *Arabidopsis thaliana* leaf epidermal cells via particle bombardment. When analysed with confocal microscopy, the fluorescence signals were obtained predominantly in plastids of transformed cells. Still, variable localization was observed at sub-plastidial level for the two constructs. In case of N-terminal 100 AA, the fluorescence signals were observed primarily in plastid stroma, while for full length TatA, fluorescent protein accumulation was observed at a sub-plastidial structure presumably the thylakoid membranes (Figure 4.2).

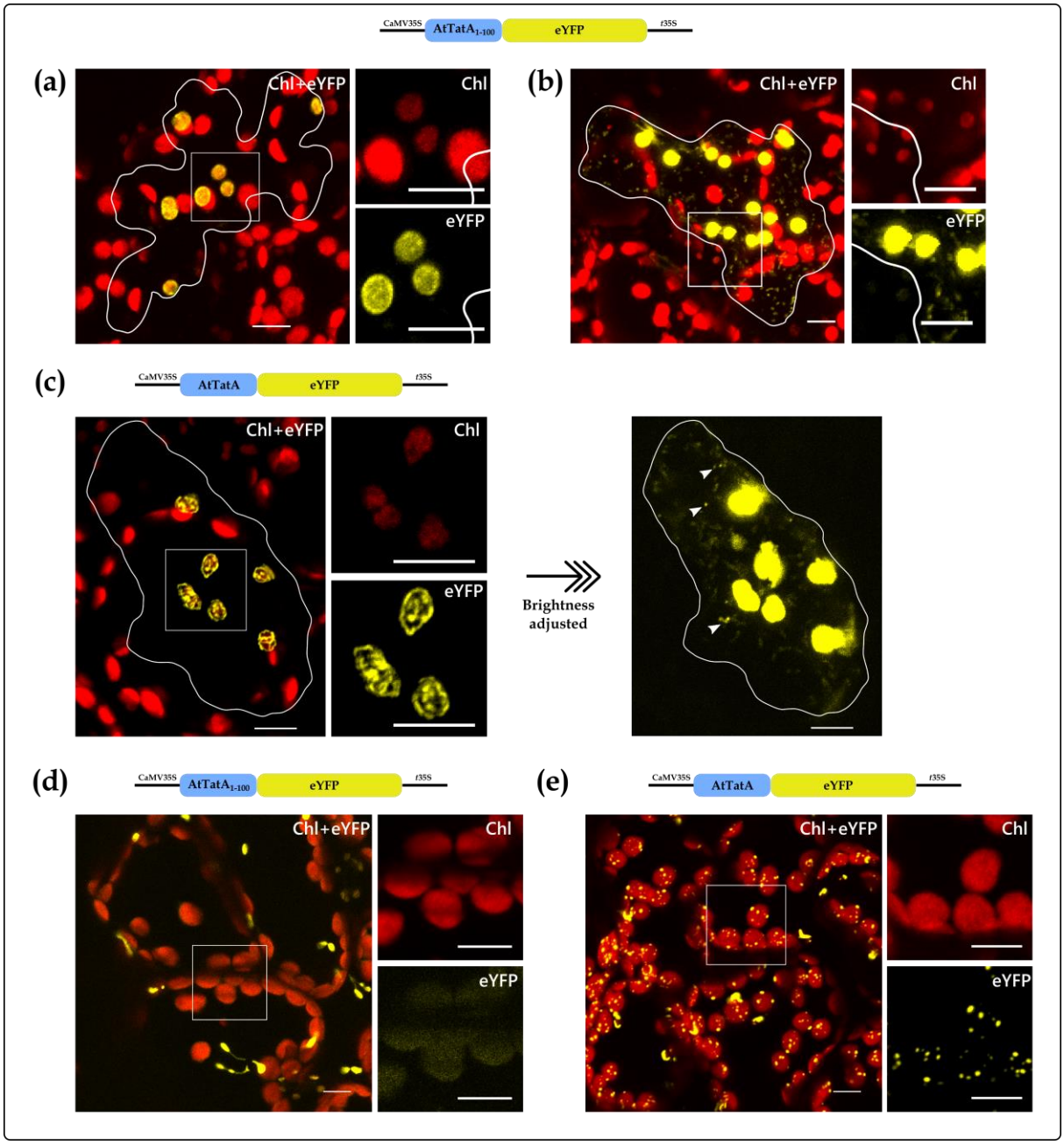


Figure 4.2 Subcellular localization of *Arabidopsis thaliana* TatA. **(a-c)** The coding sequences of *At*TatA1-100/eYFP (**a, b**) or *At*TatA/eYFP **(c)** were transiently expressed under the control of the CaMV 35S promoter after particle bombardment of leaf epidermis cells of *Arabidopsis thaliana* and analyzed by confocal laser scanning microscopy. Representative cells showing either plastid localization **(a),** or dual localization **(b)** of *At*TatA1-100/eYFP. The representative transformed cell showing **(c)** predominant thylakoid localization of *At*TatA/eYFP (*left*) while, its accumulation in punctate structures (white arrows) is visible after manual brightness adjustment (*right*). **(d-e)** The representative leaf epidermal cells of transgenic *Arabidopsis thaliana* plants constitutively expressing either AtTaA1-100/eYFP **(d)** or AtTatA/eYFP **(e)** showing exclusive plastid targeting. The punctate structures in figure **(e)** corresponds to thylakoids and should not be mistaken for mitochondria. The images are presented as overlay of the chlorophyll channel (displayed in red) and the eYFP channel (displayed in yellow). The squares highlight areas of the transformed cells that are shown in higher magnification separately for the chlorophyll channel and the eYFP channel, as indicated. The scale bars correspond to 10 µm.

Interestingly, few transformed cells (approx. 5%) expressing AtTatA$_{1-100}$/eYFP showed additional fluorescence signals in punctate structures. While, in case of AtTatA/eYFP only two such cells could be identified where punctate structures were visible after increment of brightness (Figure

4.2c). Concerning the size and morphology, these structures resembled mitochondria. This observation was striking considering the fact that other mono-specific plastid targeted transit peptides, e.g. FNR_{1-55}/eGFP, never showed such localization pattern. However, the intensity of the fluorescence signal coming from the punctate structures was significantly less compared to the plastids. Thus, to obtain a higher yield of transformed cells, the above constructions were used to transform *Nicotiana benthamiana* leaf epidermal cells via *Agrobacterium* infiltration. In line with the results from particle bombardment, the fluorescence signals were observed predominantly in plastids for both constructs (*At*Tat_{1-100}/eYFP and *At*TatA/eYFP) and additionally in punctate structures in few transformed cells in case of *At*Tat_{1-100}/eYFP. Although the fluorescence signals coming from the punctate structures were rather weak and visible when the brightness of acquired images was enhanced manually (data not shown).

Generation of stable transgenic lines of *Arabidopsis thaliana* was the next obvious step to test if the above protein localization pattern is observed in the transformed cells constitutively expressing *At*Tat_{1-100}/eYFP and *At*TatA/eYFP. Surprisingly though, the stably transformed cells of *Arabidopsis thaliana* leaf tissue did not show the above dual localization pattern. Instead, the fluorescence signals were obtained solely in plastids and the sub-plastidial protein localization was similar as observed earlier via particle bombardment, i.e., predominantly in plastid stroma with *At*Tat_{1-100}/eYFP and in thylakoid membranes with *At*TatA/eYFP (Figure 4.2d-e). This variable protein localization pattern obtained after utilization of two different experimental approaches has also been observed earlier with some previously characterized dually targeted proteins, e.g. $GrpE_{1-100}$/eYFP (see Section 2.2.3).

4.2.3 Subcellular localization of *Pisum sativum* (pea) TatA

Most mechanistic and biochemical studies exploring Tat translocase are based on pea model system. Thus, the subcellular localization of pea TatA was also examined so that the results could be subsequently confirmed via *in organello* protein import studies. For this purpose, the coding sequence for N-terminal 100 amino acids of *Pisum sativum* (pea) TatA fused to eYFP ($PeaTat_{1-100}$/eYFP) was expressed under control of CaMV35S promoter in pea leaf epidermal cells via particle bombardment. Despite having no significant similarity in primary structure of transit peptides, the $peaTatA_{1-100}$/eYFP showed similar localization behaviour as observed earlier for *Arabidopsis* $AtTatA_{1-100}$/eYFP. The fluorescence signals were observed primarily in plastids of transformed cells (Figure 4.3a) while a sub-group of transformed cells showed additional fluorescence signals in punctate structures (Figure 4.3b). In five independent particle bombardment assays, approximately 15% of cells showed such dual protein localization

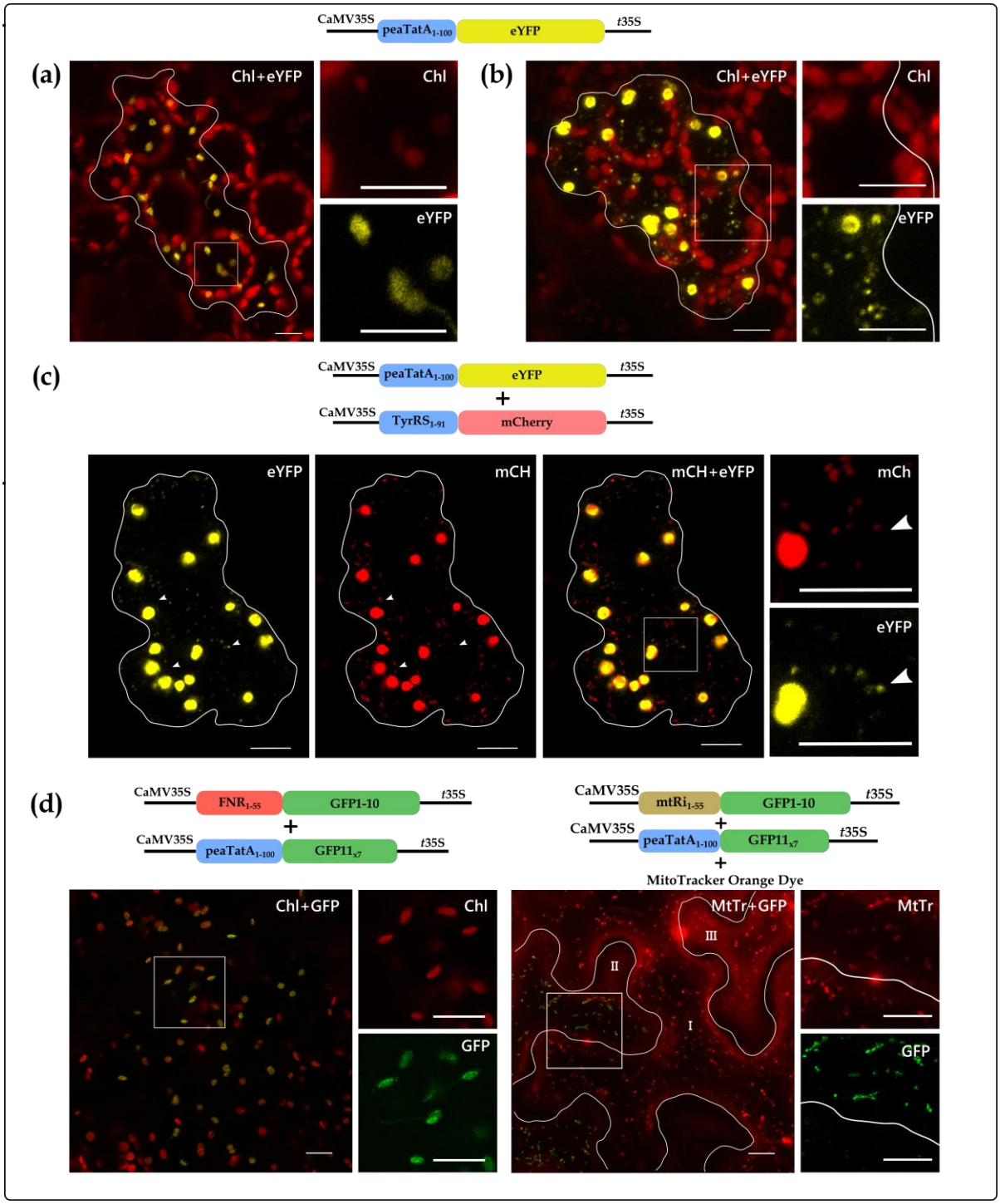


Figure 4.3 Subcellular localization of *Pisum sativum* TatA. ***(a-b)*** The coding sequence of peaTatA$_{1\text{-}100}$/eYFP was transiently expressed under the control of the CaMV 35S promoter after particle bombardment of leaf epidermis cells of *Pisum sativum* and analyzed by CLSM. Representative cells showing plastid localization of peaTatA$_{1\text{-}100}$/eYFP **(a)** or dual localization **(b)** to plastids and punctate structures (presumably mitochondria). ***(c)*** The coding sequence of peaTatA$_{1\text{-}100}$/eYFP was co-bombarded with coding sequence of AtTyrRS$_{1\text{-}91}$/mCherry on *Pisum sativum* leaf epidermal cells. A representative cell showing co-localization of eYFP with mCherry signals in plastids and partial co-localization (white arrows) of eYFP punctate structures with mCherry mitochondria signals. ***(d)*** The coding sequences of FNR$_{1\text{-}55}$/GFP1-10 and peaTatA$_{1\text{-}100}$/GFP11x7 (*left panels*) (or) mtRi$_{1\text{-}100}$/GFP1-10 and peaTatA$_{1\text{-}100}$/GFP11$_{x7}$ (*right panels*) were transiently co-expressed after *Agrobacterium* co-infiltration into the lower epidermis of *Nicotiana benthamiana* leaves and analyzed by CLSM. The representative cell showing either plastid localization of peaTatA$_{1\text{-}100}$/GFP11$_{x7}$ (left panels) or mitochondrial localization in few transformed cells (marked as II) (*right panels*). The GFP fluorescence signals emitted from mitochondria clearly co-localized with MitoTracker® Orange fluorescence signals emitted from mitochondria after infiltration of same leaf epidermis with 0.01 µM MitoTracker® Orange dye 15 min prior imaging. The scale bars correspond to 10 µm (a, b, c) and 20 µM (d).

Further, to confirm the identity of these punctate structures, co-localization analysis of peaTatA$_{1-100}$/eYFP was performed with a dual targeting peptide AtTyrRS$_{1-91}$/mCherry that labels both plastids and mitochondria. The probability of occurrence of cells, displaying dual localization and simultaneously expressing both constructs, was far less. Only, two such cells were detected after a single co-bombardment assay. In these two cells, the fluorescence signals coming from plastids in eYFP and mCherry channels showed a clear co-localization, while punctate structures visible in eYFP channel did not completely co-localize with mCherry fluorescence signals coming from mitochondria (Figure 4.3c). This observation was unexpected and due to limited number of transformed cells examined, it was difficult to draw any conclusion regarding identity of punctate structures. The existence of two subgroups of mitochondria in a single cell (Mueller et al., 2014) or potential competition between two over-expressed proteins for mitochondrial import (Langner et al., 2014) might be the possible reasons this observation.

4.2.4 Determining mitochondrial targeting specificity of *Pisum sativum* (pea) TatA with *sa*split-GFP system

The above subcellular localization assays provided a hint that peaTatA$_{1-100}$/eYFP is potentially a dually targeted protein. However, its mitochondrial targeting specificity was still unclear. Thus, to confirm that peaTatA$_{1-100}$ is indeed targeted to mitochondria, a more sensitive self-assembling split-GFP (*sa*split-GFP) based approach was utilized (see Chapter 3). The construct, carrying gene sequence coding for peaTatA$_{1-100}$ fused to a heptad of GFP11-repeats (peaTatA$_{1-100}$/GFP11$_{x7}$), was co-infiltrated with either plastid targeted FNR$_{1-55}$/GFP1-10 or mitochondria targeted mtRi$_{1-100}$/GFP1-10 into lower epidermis of *Nicotiana benthamiana* leaf. The cells co-transformed with peaTatA$_{1-100}$/GFP11$_{x7}$ and FNR$_{1-55}$/GFP1-10 (targeted to plastids) showed fluorescence signals exclusively in plastids (Figure 4.3d, *left panel*). On the other hand, majority of cells co-transformed with peaTatA$_{1-100}$/GFP11$_{x7}$ and mitochondria targeted mtRi$_{1-100}$/GFP1-10 showed no fluorescence signals. Again, few isolated cells showed fluorescence signals in punctate structures as observed earlier with particle bombardment (Figure 4.3d cell II, *right panel*). To further clarify, if the punctate fluorescence structures are indeed mitochondria the transformed leaf area was infiltrated with 0.01 µM solution of fluorescent MitoTracker® Orange dye to label mitochondria in leaf epidermal cells. The fluorescence signal coming from MitoTracker channels showed clear co-localization with green fluorescence signals confirming the mitochondrial targeting specificity of peaTatA$_{1-100}$/GFP11$_{x7}$ (Figure 4.3d *right panel*).

4.2.5 Generation of TatA *knock-out* mutant of *Arabidopsis thaliana*

Above experiments, demonstrate the potential dual targeting of TatA to plastid and mitochondria. However it was still a question if the mitochondrial targeting of TatA is indeed functionally

relevant? One possible way to answer this question could be functional characterization of TatA *knock-out* mutants. However, these mutants are neither available in public repository nor have been characterized for any plant species so far. Thus, the TatA *knock-out* lines were generated using Crispr/Cas9 based genomic deletion approach. In *Arabidopsis thaliana* genome, the *AtTatA* gene is present in a single copy and this model system can be easily transformed. Hence, this model plant was selected for generation of *knock-out* mutant.

For this purpose, two sgRNAs targeted to two distant regions (1397 bp apart) of *AtTatA* gene were utilized as described in Ordon et al. (2016) (see Figure 4.4a). *Arabidopsis thaliana* accession Columbia-0 (Col-0) plant lines were transformed with the construct carrying genes, coding for hCas9 protein and a gRNA (guide RNA). The transgenic plants were chosen by polymerase chain reaction (PCR) based genotyping approach with *AtTatA* gene specific primers. Later, the next generation plants carrying a deletion in *AtTatA* gene and lacking the transgenes, i.e. hCas9 and gRNA, were selected and propagated for further analysis. PCR amplification and sequencing of gene product from *tatA* mutant lines revealed a 658 bp deletion within the genomic region spanning *TatA* gene (Figure 4.4c). The homozygous *tatA* knock out plants had white cotyledons and were seedling lethal, i.e. were unable to survive after 10-12 days of germination. In contrast to that, heterozygous plants showed phenotype similar to wild-type Col-0 plants (Figure 4.4b). The immunoblot analysis of total protein extract from mutant seedlings confirmed the absence of TatA protein from these mutant lines (Figure 4.4d).

To further validate that this seedling lethal phenotype is due to lack of TatA protein, the heterozygous *tatA* (+/-) plants were transformed with construct encoding full-length *At*TatA fused to eYFP expressed under control of CaMV35S promoter. The transformed plants were first selected on antibiotic containing growth medium and later screened for a homozygous *tatA* (-/-) background via genotyping. The expression of chimeric AtTatA/eYFP was able to complement the *tatA* (-/-) mutant phenotype (Figure 4.5c), demonstrating that the visible phenotype was indeed due to the absence of TatA in the mutant lines. Similarly, the mutant plants were transformed with a construct carrying *AtTatA* gene (approx. 3.4 kb genomic DNA fragment) with its untranslated regions and regulatory elements. These transgenic homozygous *tatA* (-/-) plants expressing native TatA were also able to complement the mutant phenotype (data not shown).

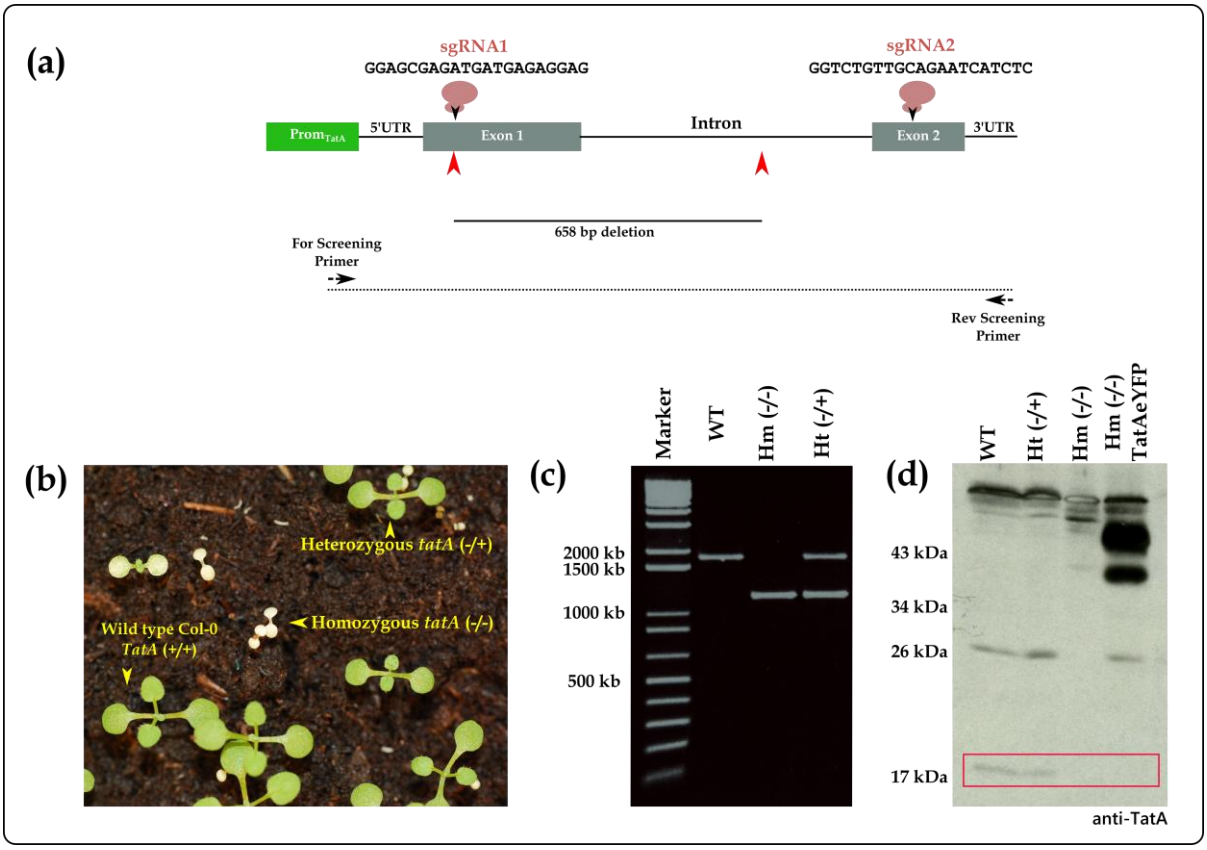


Figure 4.4 Generation of TatA *knock-out* mutant of *Arabidopsis thaliana*. **(a)** Gene map of *Arabidopsis thaliana* TatA indicating the guide RNA recognition sites (black) and sites of actual deletion (red) **(b)** The seeds from T3 generation heterozygous mutant plants were germinated on soil and the seedlings were confirmed for their zygosity via PCR with *TatA* gene specific primers. **(c)** A representative PCR amplification of TatA gene from genomic DNA templates of wild type, homozygous *tatA* (-/-) or heterozygous *tatA* (-/+) seedlings. **(d)** Western blot analysis of total protein extracts from leaf tissue of wild type, homozygous *tatA* (-/-), heterozygous *tatA* (-/+) and complemented TatA lines with anti-TatA antibodies (performed by Dr. Daniela Rödel) indicate absence of 17 kDa TatA protein fragment from *tatA* mutant plants.

4.2.6 Ultrastructure of subcellular compartments in TatA *knock-out* mutant

The white cotyledons and seedling lethal phenotype of *tatA* provided a hint that the plastid metabolism might be altered in these mutant lines. Thus, in order to examine the effect of *tatA* deletion on the cellular ultrastructure and organelle morphology, the electron microscopy of cotyledon tissue sections was performed (by Dr. Gerd Hause at Biozentrum, MLU Halle). Specimens were prepared after sectioning, staining and mounting of cotyledons from 10 days old seedlings of *tatA* mutants, complemented transgenic *tatA* lines carrying *At*TatA/eYFP and wild type Col-0 plants. The electron micrographs revealed that the chloroplasts of *tatA* mutants were devoid of internal membrane structures and were highly vacuolated (Figure 4.5b). The altered chloroplast morphology in *tatA* mutants was similar to what has been observed earlier for *tatB* and *tatC T-DNA* mutants of *Arabidopsis thaliana* (Motohashi et al., 2001; Wang et al., 2016), indicating the absolute functional requirement of all three subunits of Tat translocase in *Arabidopsis thaliana*. On the other hand, the chloroplast ultrastructure of complemented

transgenic mutant lines carrying *At*TatA/eYFP were similar to wild type Col-0 lines indicating the re-constitution of Tat translocase in chloroplasts with chimeric TatA/eYFP.

Surprisingly, the mitochondrial ultrastructure in *tatA* mutants remained unaffected. Instead, the mitochondria of complemented *tatA* lines, carrying *At*TatA/eYFP, were vacuolated and devoid of internal structure (Figure 4.5c). However, in the complemented lines, no eYFP fluorescence was observed in mitochondria when analyzed via fluorescence microscopy. Thus, it was a bit difficult to conclude if the observed mitochondrial morphology is merely due to overexpression of chimeric *At*TatA/eYFP or due to its targeting to mitochondria and perturbation of putative Tat machinery. Furthermore, the electron micrographs corresponds to a single transgenic line and the observed alterations in mitochondrial ultrastructure could be an artefact of *T-DNA* integration and/or counter selection.

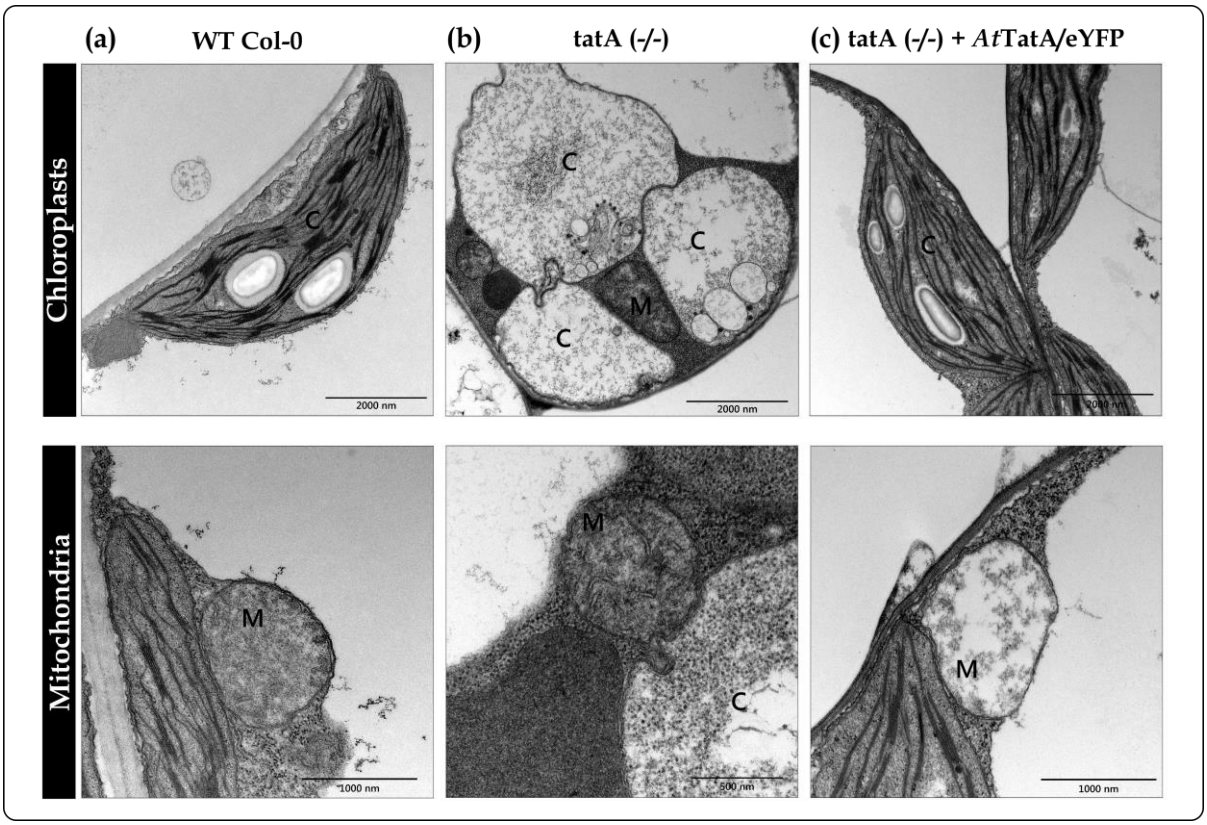


Figure 4.5- Electron micrographs of Arabidopsis thaliana *tatA* mutant. The ultrastructure of chloroplast (upper panels) and mitochondria (lower panels) from two weeks old cotyledons of (a) wild-type Arabidopsis thaliana, (b) *tatA* mutant line, and (c) *tatA* mutant complemented with *At*TatA/eYFP. The micrograph suggests that the chloroplast were highly vacuolated and their morphology was altered in tatA mutants while mitochondrial ultrastructure was altered in complemented lines. The respective organelles are depicted a C (chloroplast) or M (mitochondria). Electron microscopy was performed in collaboration with Dr. Gerd Hause (at Biozentrum, MLU Halle). Scale bar as indicated.

4.3 Discussion

Present study proposes that TatA retains dual targeting specificity to both mitochondria and plastids. However, it was tedious to determine mitochondrial targeting specificity of TatA using fluorescence protein (FP)-tagging approaches. The results indicated that TatA is preferentially targeted to plastids while its mitochondria targeting was comparatively weaker and even remained undetected in most of the transformed cell. A lower demand of TatA in mitochondria might be one of the potential reasons for this observation. In plant mitochondria, mtRi seems to be the only potential substrate for Tat translocase. Thus, only few molecules of TatA might be sufficient for efficient translocation of a single substrate (e.g. Hauer et al., 2013).

Besides this, a high turnover rate of FP-tagged TatA protein or potential degradation by quality control system in the mitochondria might be another reason for variable results obtained after transient and stable transformation. In stable transgenic lines, the timescale between expression of protein and detection via microscopy is comparatively longer and sufficient either for counter-selection by plants or activation of specific protein degradation pathway. Thus, a protein with high turnover rate inside one of the organelle cannot be detected in the transgenic lines. While in case of transient assays, the subcellular localization of a protein can be detected as short as within first 16 h of its expression.

In a similar fashion, the fluorescence microscopy based *in vivo* approaches failed to determine the subcellular localization of TatB-like protein. When analyzed via two different transient transformation assays FP-tagged TatB-like protein localized only to cytosolic aggregates. On the other hand, the complementary *in vitro* import assays clearly demonstrated the targeting of authentic precursor peaTatA as well as AtTatB-like protein in isolated mitochondria from pea leaf tissues (Bennewitz et al., unpublished data) suggesting that mitochondrial translocation machinery could in principle recognize and facilitate import of both proteins. Similarly, a recent mitochondria complexome of *Arabidopsis thaliana* indicated the existence of TatB-like protein in an approx. 600 kDa protein complex (Senkler et al., 2017). However, the other two proteins, i.e. TatA and mtTatC remained undetected in this complex.

The expression of mtTatC transcripts was shown to be up-regulated up to 6 fold upon salicylic acid treatment (van der Merwe and Dubery, 2007). Similarly, the stress conditions are known to affect the import of certain mitochondrial protein (Taylor et al., 2003). This raises a possibility that mitochondria recruit the Tat machinery under stress conditions to transport specific substrates. Thus, differential expression of TatC substrate and variable import of TatA proteins might be another potential reason why these proteins remained undetected in majority of

mitochondrial proteomes. To further clarify the existence of functional TatA and a functional Tat machinery in plant mitochondria, experiments including subcellular localization under stress conditions, immunoblots of organelle protein fraction isolated after stress treatment or immunogold labelling of leaf tissue with Tat-specific antibodies could be helpful.

Chapter 5-
General Discussion

Protein transport across the membrane is an outcome of cooperation between membrane transport apparatus and protein transport signal. The first-hand interaction between the two decides the fate of protein transport into a membrane-bound cell organelle. Once transported, depending on the function, the proteins either accumulate or reside transiently in a particular organelle. The extent of accumulation is variable for different proteins and might even spatiotemporally regulated in the cell. For example, the dual localization of, otherwise plastid targeted, accelerated cell death 2 (ACD2) protein of *Arabidopsis thaliana* is evident only upon pathogen infection (Yao and Greenberg, 2006). This could possibly explain why none of the experimental approaches alone, including high throughput biochemical assays, could identify the complete set of proteins in a specific subcellular compartment. For example, in *Arabidopsis thaliana* total number of predicted proteins in plastids is 2000-3000 while only approx. 1,200 have been experimentally verified as plastid proteins (van Wijk and Baginsky, 2011). Similarly, more than 2000 proteins are predicted to reside in mitochondria of higher plants and only approx. 1000 of these proteins were detected in a large-scale proteomic analysis (Rao et al., 2017).

5.1 Experimental approaches to define dual protein targeting

In this study, a comparison of different "established" experimental approaches also suggest that a single method is not sufficient to determine targeting specificity of a protein in an organelle and each method can only address particular aspects of the entire sorting and transport process (Figure 5.1). For example, while *in organello* experiments allow to analyze authentic precursor proteins without any heterologous reporter fusion, they can only answer the question if the protein import machinery of a given organelle is principally able to recognize this precursor as substrate and translocate it across the organelle envelope membranes. Whether this will in fact take place also in an intact plant cell must remain unsolved due to the lack of cytosolic targeting factors in such assays. Similarly, since the purified organelles are a pre-requisite in these assays, consequently loss of some specific translocase receptors on outer membrane during the process of purification could led to deviating results. On the other hand, all *in vivo* approaches, which allow to analyze protein transport in the natural context of the cell, depend on chimeric reporter constructs, which in turn might affect the actual membrane translocation process, for example due to their particular folding properties (e.g., Marques et al., 2004). Furthermore, it is not always

possible in such assays to distinguish between mere binding of the reporter construct to the surface of the organelle and its actual import (Baudisch and Klösgen, 2012). This is particularly difficult in those instances when import into small organelles like mitochondria is studied. Finally, transgenic plants expressing such reporter constructs have the additional problem of potential counter selection or accumulation of mutations, which might take place during the long-lasting regeneration process after transformation. Such processes might lead to silencing of transgene expression or even growth inhibition of transformants expressing the transgene. It is evident that all this could significantly distort the actual targeting result.

In this study, some of these methodological constraints were addressed by modifying reporter proteins system with the use of *sa*split-GFP technology (Cabantous et al., 2005). Application of this technology significantly reduces the reporter protein 'burden' on candidates during *in vivo* protein localization assays, which otherwise could affect the membrane protein transport. Similarly, the *sa*split-GFP system made it possible to determine precise sub-organellar location of a candidate protein without the use of ultra-modern microscopy devices. Separation of fluorescence signals for dually targeted proteins in each of the target organelle opened up ways for quantitative cell biology to study spatial and temporal dynamics of dual protein import into mitochondria and plastids. However, one has to be aware that, despite all these improvements, there is still no general "golden standard" available when analyzing the specificity of protein targeting in eukaryotic cells. Instead, the obvious consequence must be to employ at least two principally different methods in such analyses to achieve reliable conclusions. Combining different experimental approaches could be the best-case scenario to assess the targeting specificity of a protein with high accuracy (see Table 5.1 to figure out potential of each experimental approach analyzed in this study).

5.2 Evolutionary significance of dual protein targeting

Irrespective of a few potential false positives resulting from the general technical limitations addressed above, the current estimate that approximately 5% of total organellar proteins are probably dually targeted to mitochondria and chloroplasts (Mitschke et al., 2009; Baudisch et al., 2014) is remarkably high and exceeds by far those assumed for any other pair of organelles. Technically speaking, the main reason is the considerable similarity of the mitochondrial and plastidial transit peptides, which are difficult to distinguish even for prediction programs like TargetP. In fact, it was recently shown that the organelle specificity of these targeting signals can easily be modulated towards mitochondria by the incorporation of individual arginine residues into the N-terminal regions which led to the hypothesis of a positively charged "avoidance signal"

for chloroplast import (Ge et al., 2014). In the same line, positive and negative charges in such transit peptides, derived from either amino acid residues or phosphorylation, were proposed to provide an important driving force in the evolution of organelle transport specificity (Garg and Gould, 2016). And finally, comparison of mitochondrial transit peptides from plant and non-plant species indicated that these transit peptides had been adapted after the establishment of chloroplasts as the second endosymbiotic organelle to achieve higher organelle specificity (Staiger et al., 2009).

Table 5.1 Potential of different experimental approaches to determine various aspects of protein targeting into plant cell organelles. *Green* boxes indicate the approaches most suitable to study particular aspect.

	in vitro	*in vivo*					
		Fluorescent protein tagging				Alternative Approaches	
Aspects of protein Targeting	***in organello* import**	**Particle Bombardment**	**Protoplast Transformation**	**Agro-infiltration**	**Transgenic plants**	***sa*split-GFP system***	**Phenotype Complementation**
Targeting specificity to organelles	+++	++	++	++	+	+++	+++
Targeting specificity in a cellular context	-	++	+	+++	++	++	+++
Targeting specificity in a homologous system	-	+++	+++	-	+	+	-
Transient targeting to an organelle	++	++	++	++	-	+++	-
Protein accumulation	++	+	+	++	++	++	+
Tissue specific targeting	-	-	+	-	+++	++	-
Development/stress specific targeting	-	-	-	-	+++	++	-

*(+++) Highly suitable; (++) Suitable; (+) Maybe optimal for some proteins; (-) not suitable. *The sasplit-GFP system can be used in combination with several different* in vivo *approaches.*

In a wider context, the phenomenon of frequent dual protein targeting into both endosymbiotic organelles appears to be an immediate consequence of the two consecutive evolutionary events leading to the establishment of mitochondria and chloroplasts (Lang et al., 1999; McFadden, 1999). In the first event, the mitochondrial transport signals and protein import machineries had probably to be developed from scratch, because to the best of our knowledge there is no case study of free-living bacteria performing import of proteins carrying specific transport signals. In consequence, neither the transport signals nor the import machineries could have been particularly selective at first. Otherwise, organelle transport of those proteins that were derived from genes transferred to the nucleus early in evolution could not have taken place. In contrast,

during the second endosymbiotic event, when the prokaryotic chloroplast ancestor was engulfed, both transport signals and import machinery had already been established for mitochondria. Hence, for the chloroplast a model for protein translocation system was available by means of duplication of genetic information and subsequent modification. Selective pressure presumably then modulated these transport signals to become more organelle-specific. As a result, for those proteins that have a negative impact on the metabolism of the "wrong" organelle, strictly monospecific transit peptides are found today, whereas in the other instances a certain degree of "mistargeting" is apparently tolerated by the cell. This might explain the remarkable similarity of transit peptides for mitochondria and chloroplasts and suggests that dual targeting of proteins was probably the rule rather than exception immediately after establishment of chloroplasts.

5.3 Functional relevance of dual protein targeting

The evolutionary aspects of dual protein targeting could also explain why for only a relatively small proportion of dually targeted proteins a function in both mitochondria and plastids can be assumed. These include, for example, aminoacyl-tRNA synthetases (Duchêne et al., 2005), RNA polymerase RpoTmp (Hedtke et al., 2000; 2002), the ribosomal protein S16 (Ueda et al., 2008), and, as recent additions to this list, presequence protease AtPreP (Kmiec et al., 2014), methyltransferase PrmA (Mazzoleni et al., 2015), and pentatricopeptide repeat (PPR) proteins (Guillaumot et al., 2017). In some cases, even if functional, the proteins may not be essential for both organelles as shown for RNA polymerase RpoTmp (Tarasenko et al., 2016).

For the majority of proteins that show dual targeting in the experimental approaches described in this study, a function in both endosymbiotic organelles appears even unlikely. For example, the mitochondrial proteins GCS-H, a component of the glycine decarboxylase system, is unlikely to play an active role in chloroplast metabolism, despite the fact that for this protein dual targeting was observed in all experimental approaches (Figure 2.2). Similarly, the mitochondrial targeting specificity of RbcS (Rubisco small subunit) does not appears to be functionally relevant. It almost seems as if such ‘mistargeting’ of a considerable number of proteins is not only acceptable for the cell but might even serve as a kind of "evolutionary playground". In line with that, it was speculated that such frequent dual targeting might lead to a positive selection pressure and could make possible the transfer of complete metabolic pathways across organelle borders or development of novel pathways in organelle (Martin, 2010). The latter seems particularly true with regards to the evolution of potential Tat machinery in mitochondria.

5.4 Perspectives

The chloroplastic TatA presents one such hypothetical scenario where its dual targeting might have supported the evolution of a Tat translocase in mitochondria. However, its functional relevance in mitochondria is yet to be experimentally verified. Multiple experiments can be designed to delineate the function-location nexus when the study is focused on a single protein. For example, characterization of mutant lacking the candidate protein and its complementation with similar protein but fused to different (monospecific) transit peptides could give a hint about its function (e.g., Xu et al., 2012 and Kmiec et al., 2014). Likewise, protein localization studies under variable growth conditions or in a tissue-specific manner could be useful in understanding spatiotemporal regulation of dual protein targeting (e.g., Engel et al., 2011). On the other hand, it sounds rather impractical to apply all these experimental approaches when analyzing a large set of proteins. This could possibly explain why the phenomenon of dual targeting is still naïve. The list of such dually targeted proteins is expanding with the experimental advancements in the field while for majority of them a function remains enigmatic in one or the organelle.

Chapter 6-
Materials and Methods

6.1 Materials

6.1.1 Chemicals

Most of the chemicals utilized in this study, unless otherwise stated, are obtained from these companies: Difco-Laboratories (Detroit, USA), Duchefa Biochemie (Haarlem, Netherlands), Merck (Darmstadt, Germany), Roth (Karlsruhe, Germany), Serva (Heidelberg, Germany) and Sigma-Aldrich (München, Germany).

6.1.2 Antibiotics

Table 6.1 List of Antibiotics

Antibiotic	Working conc.	Solvent	Company
Ampicillin	100 µg/ml	H_2O	SERVA
Carbenicillin	100 µg/ml	H_2O	SERVA
Gentamycin	25 µg/ml	H_2O	SERVA
Hygromycin-B	30 µg/ml	H_2O	Duchefa Bioch.
Kanamycin	50 µg/ml	H_2O	SERVA
Spectinomycin	100 µg/ml	H_2O	Duchefa Bioch.
Rifampicin	100 µg/ml	DMSO	Sigma-Aldrich

6.1.3 Enzymes

All restriction enzymes (five-buffer system) were purchased from Thermo Fisher Scientific (Massachusetts, United States). The following enzymes were utilized and purchased from mentioned company.

Table 6.2 List of enzymes

Enzyme	Company
T4 DNA ligase	Thermo Fisher Scientific
FastAP™ Thermosensitive Alkaline Phosphatase	Thermo Fisher Scientific
NZYTaq DNA polymerase	NZYTech Lisboa, Portugal
Biozym Taq DNA Polymerase	Biozym Scientific GmbH, Germany
Phusion® High-Fidelity DNA Polymerase	Thermo Fisher Scientific
Macerozyme R-10	Duchefa Bioch. (Haarlem, Netherlands)
Cellulase R-10	Duchefa Bioch. (Haarlem, Netherlands)

6.1.4 Molecular Biology Kits

(a) NucleoSpin Plasmid - Plasmid Miniprep kit, **(b)** NucleoBond® PC 100 Plasmid Midiprep kit and **(c)** NucleoSpin Gel and PCR Clean-up kit were obtained from Macherey-Nagel GmbH & Co. KG (Schkeuditz, Germany). The **(d)** CloneJET™ PCR Cloning Kit was obtained from Thermo Fisher Scientific (Massachusetts, United States). **(e)** MoClo Plant Parts Kit I and II (Weber et al., 2011; Engler et al., 2014; Gantner et al., 2018)

6.1.5 Vectors (Molecular Biology)

Table 6.3 Parental vector backbones utilized in this study. (See section 6.2.1 for a list of newly generated vectors.)

Name	Vector type	Bacterial Selection*	Plant Selection	Reference
pRT100mod (Ω/Not/Asc)	Cloning	Amp	-	Überlacker and Werr (1996)
pCB302	Binary	Kan	Glufosinate	Xiang et al. (1999)
pLSU4GG	Binary	Kan	Hyg^+	Erickson et al. (2017)
MoClo Plant Parts I and Plant Parts II	Multiple vectors	Misc.	None	Weber et al. (2011), Engler et al. (2014), Gantner et al. (2018)

*Amp- Ampicillin, Kan- Kanamycin

Table 6.4 Vectors obtained from mentioned sources.

Name	Bacterial Selection*	Reference
pRT100mod:MCS:eYFP	Amp	Baudisch et al. (2014)
pRT100mod:GCS100eYFP	Amp	
pRT100mod:Ef-Tu100eYFP	Amp	
pRT100mod:GrpE100eYFP	Amp	
pRT100mod:PDF100eYFP	Amp	
pRT100mod:mtRi100eYFP	Amp	Rödiger et al., 2011
pRT100mod:FNR55eGFP	Amp	Rödiger et al., 2011
pLSU4GG:FNR55eGFP	Kan	Erickson et al. (2017)
pCB302:mtRi100eYFP	Kan	Rödiger et al., 2011
pRT100mod:peaTatA100eYFP	Amp	Available in Klösgen's Lab
pRT100mod:AtTatA100eYFP	Amp	

*Amp- Ampicillin, Kan- Kanamycin

6.1.6 Oligonucleotides (Oligos)

The oligonucleotides were synthesized by metabion international AG (Planegg, Germany)

Table 6.5 List of oligonucleotides

Oligos	Sequence (5' to 3')
Primers for Golden Gate cloning from pRT100 to pLSU4GG	
pRT100 GG Forward	ATATGGTCTCTAGGTAACATGGTGGAGCACGACACTCTCG
pRT100 GG Reverse	ATATGGTCTCTAAGCGCAGGTCACTGGATTTTGGTTTTAGG
Sequencing Primers	
CaMV Forward	ATGACGCACAATCCCACTATCC
CaMV Reverse	TGCTCAACACATGAGCGAAACC
pLSU4GG Seq Reverse	CTCGAACAAGACGTTTCCC

Oligos	Sequence (5' to 3')
GFP1-10 Seq Rev	CTGAACTCCGTACGTAAGCG
GFP11 Seq Rev	CAATGAAACTGATGCATTG
	Primers for PTOX cloning in pRT100mod vector
PTOX For (NcoI)	AGACCATGGCGGCGATTTCAGGCATC
PTOX Rev (BsrGI)	AGTTGTACATTAACTTGTAATGGATTTCTTGAGG
PTOX mat For (NcoI)	AGACCATGGCAACGATTTTGCAAGACGAT
	Primers for *TatA* cloning
AtTatA Forward	GGATCCTATGGCGACATCTGTTGCG
AtTatA Reverse	GCCATGGCTACATTCTCCTTTGAGCTTGAAG
	Primers for *tatA* mutants generation and screening
AtTatA Scr Forward	AAAGCCCATTAGAACTAGGTTGGATC
AtTatA Scr Reverse	GTAGCTACAAGGAAAAGCAAACACT
hCas9 Scr For	AAAAGGGTATAAAAGAACTGGGGTC
hCas9 Scr Rev	GATTTCTCCTGTTTCTCCGTTTGTT
CrsprOligo_TatA_F_3	ATTGGAGCGAGATGATGAGAGGAG
CrsprOligo_TatA_R_3	AAACCTCCTCTCATCATCTCGCTC
CrsprOligo_TatA_F_16	ATTGGTCTGTTGCAGAATCATCTC
CrsprOligo_TatA_R_16	AAACGAGATGATTCTGCAACAGAC
	Primers for *sa*split-GFP cloning in pRT100mod
GFP1-10 For (BsaI)	TATGGTCTCTCATGTCCAAAGGAGAAGAACTGTTTA
GFP1-10 Rev (XbaI)	CCTCTAGACTAACTTCCGCCGCCACCT
GFP11 For (NcoI)	ATCCATGGGCGGCAAATTCATGCGTGACCACATGGT
GFP11x7 Rev (XbaI)	ATTCTAGATTAGGTGATACCGGCAGCATTGAC
GFP11x3 Rev (XbaI)	ATCTAGATTATCCGGTTATTCCGGCTGCATT
GFP11x1 (XbaI)	ATCTAGATTATGTAATCCCAGCAGCATTTAC
	Primers for level 0 of PlaMiNGo toolkit
FNR55 For (pTEI119)	TTTGAAGACATCCATGACCACCGCTGTCACC
FNR55 Rev (pTEI119)	TTTGAAGACTACATTGCGGCCCTGATGGGTCCCATTTTC
mtRi100 For (pTEI120)	TTTGAAGACATCCATGCTTCGAGTAGCAGGTAG
mtRi100 Rev (pTEI120)	TTTGAAGACATCATTGCGCTAGGATCTCCAGGTGGAT
GFP1-10 For (pTEI121)	TTTGAAGACATAATGTCCAAAGGAGAAGAAC
GFP1-10 Rev (pTEI121)	TTTGAAGACTAAAGCTAACTTCCGCCGCCACCTG
ccdB For (pTEI122)	TTTGAAGACATAATGTGAGACCGACTGGCTGTGTATAAGGG
ccdB Rev (pTEI122)	TTTGAAGACTACGAATGAGACCTTGATCGGCACGTAAGAGG
GFP11x3 For (pTEI123)	TTTGAAGACATTTCGATGGGCGGCAAATTCATGCG
GFP11x3 Rev (pTEI123)	TTTGAAGACTAAAGCTTATCCGGTTATTCCGGC
GFP11x7 For (pTEI124)	TTTGAAGACATTTCGATGGGCGGCAAATTCATGC
GFP11x7 Rev (pTEI124)	TTTGAAGACTAGATCTCGTCCGCCCGACCC
GFP11x1 For	TTCGATGGGCGGCAAATTCATGCGTGACCACATGGTCCTTCATGAGTATGTAAA TGCTGCTGGGATTACATAA

Oligos	Sequence (5' to 3')
GFP11x1 Rev	AAGCTTATGTAATCCCAGCAGCATTTACATACTCATGAAGGACCATGTGGTCAC GCATGAATTTGCCGCCCAT

Primers for cloning of candidate proteins in pLaMiNGo vectors (With BsaI recognition seq)

TyrRS91 For	TTTGGTCTCTAATGGCATATGCAACAGGAA
TyrRS91 Rev	TTTGGTCTCACGAAGCGGTTTCTTCAAGTATATCG
FNR55 For	TTTGGTCTCAAATGACCACCGCTGTCACC
FNR55 Rev	TTTGGTCTCACGAAGCCCTGATGGGTCCCATTTTC
mtRi100 For	TTTGGTCTCAAATGCTTCGAGTAGCAGGTAG
mtRi100 Rev	TTTGGTCTCACGAACTAGGATCTCCAGGTGGAT
GCS100 For	TTTGGTCTCAAATGGCACTAAGAATGTGGGCT
GCS100 Rev	TTTGGTCTCACGAACTTGTTGCCTTCACACTCTC
Gtred100 For	TTTGGTCTCAAATGGCGGCTTCTTTATCG
Gtred100 Rev	TTTGGTCTCACGAAACTCTGACGGCTAGTGAGCT
GAPDH100 For	TTTGGTCTCAAATGGCCACACATGCAGCTC
GAPDH100 Rev	TTTGGTCTCACGAAGCCCAACACCTAAGAAAGTTC
ATPS100 For	TTTGGTCTCAAATGGCGAGTCGGCGAATCT
ATPS100 Rev	TTTGGTCTCACGAAGCTCTCACATCAACAATGGCA
CoxIV29 For	TTTGGTCTCAAATGCTTTCACTACGTCAA
CoxIV29 Rev	TTTGGTCTCACGAAGGTTTTTGCTGAAGCAG
SSU79 For	TTTGGTCTCAAATGGCTTCCTCAGTTCTTTC
SSU79 Rev	TTTGGTCTCACGAACTCAAATCAGGAAGGTAT
peaTatA100 For	TTTGGTCTCAAATGGAGATAACACTTTCC
peaTatA100 rev	TTTGGTCTCACGAAGCTGCTGCCTGTTGAAAGCT

*Oligonucleotides for RF cloning are not mentioned in the list. They were designed according to Bond and Naus (2012)

6.1.7 Miscellaneous

(a) The PTOX specific antibodies were obtained from Agrisera (Vännäs, Sweden). (b) For immunoblotting, the 0.2 µM PVDF (Amersham Hybond) membranes were obtained from GE Healthcare Life Sciences. (c) Thermo Page Ruler (Prestained protein ladder) and (d) 1 Kb Plus DNA ladder were purchased from Thermo Fisher Scientific. (e) Murashige and Skoog medium (with vitamins) was obtained from Duchefa Biochemie (Haarlem, Netherlands). (f) MitoTraker Oragne® CMTMRos dye was obtained from Molecular Probes (Oregon, USA)

6.1.8 Biological Material

Plants: (a) *Arabidopsis thaliana* eco. Columbia-0 (b) *Arabidopsis thaliana immutans* mutants (provided by Prof. Steven Rodermel, ISU, United States) (c) *Nicotiana benthamiana* and (d) *Pisum sativum* var. Feltham First

Bacterial Strains: (a) *Escherichia coli* DH5α (Invitrogen), (b) *Escherichia coli* DB3.1 *ccd*B survival cells (Thermo Fisher Scientific), (c) *Agrobacterium tumefaciens* GV3101 (pMP90) (Koncz and Schell. 1986).

6.1.9 Culture Media

Table 6.6 List of culture media

Growth Medium	Organisms	Content
LB (Luria-Bertani) liquid medium	Bacteria (*E.coli*, *Agrobacterium*)	1% w/v NaCl, 1% w/v Peptone, 0.5% w/v Yeast Extract, pH 7.0
LB (Luria-Bertani) solid medium	Bacteria (*E.coli*, *Agrobacterium*)	1% w/v NaCl, 1% w/v Peptone, 0.5% w/v Yeast Extract, pH 7.0 and 1.2% Agar
YEBS Medium	*Agrobacterium*	0.5% w/v Beef extract, 0.5% w/v Sucrose, 0.5% w/v Peptone, 0.1% w/v Yeast extract, 0.1% w/v $MgSO_4.7H_2O$, pH 7.0
MS Agar Medium	*Arabidopsis thaliana*	4.4 g/l MS (with Vitamins), 0.5 g/l MES, pH 5.7 and 0.8% MicroAgar (Duchefa Biochemie)
½ MS Agar Medium	*Arabidopsis thaliana*	2.2 g/l MS (with Vitamins), MES 0.5 g/l, pH 5.7 and 0.8% MicroAgar (Duchefa Biochemie)

6.2 Methods

6.2.1 Molecular Cloning

(a) eYFP reporter constructions (Chapter 2)

The eYFP reporter constructs used for particle bombardment and protoplast transformations are based on vector pRT100mod and described in Baudisch et al. (2014) (Tabel 6.4). For subsequent cloning into a binary vector pCB302 (Xiang et al., 1999), the entire gene constructs (encoding GCS_{1-100}/eYFP, $GrpE_{1-100}$/eYFP, EF-Tu_{1-100}/eYFP, and $mtRi_{1-100}$/eYFP) with CaMV35S promoter and CaMV35S terminator were recovered after digestion with SdaI and ligated into linear pCB302 (digested with PstI and subsequently dephosphorylated). Reporter constructs PDF_{1-100}/eYFP and FNR_{1-55}/eGFP were instead inserted into the binary vector pLSU4GG (Appendix II) (Erickson et al., 2017) using 'standard' BsaI restriction/ligation Golden Gate reaction (see section 6.2.1g). The constructed binary vectors were utilized for *Agrobacterium* infiltration of *N. benthamiana* and floral dip transformation of *Arabidopsis thaliana*.

(b) Generation of sasplit-GFP vectors (Chapter 3)

The GFP1-10 and $GFP11_{x7}$ fragments were amplified by PCR from plasmids pcDNA3.1-GFP (1-10) and pACUH-$GFP11_{x7}$-mCherry-ß-tubulin (a gift from Bo Huang lab; Addgene#70218 and #70219) and cloned into pRT100mod-based vectors (Baudisch et al. 2014) either with digestion/ligation or with Restriction Free cloning (Bond and Naus, 2012). The gene sequence

coding for the N-terminal 91 amino acids of TyrRS was amplified from a vector provided by E. Glaser (SU, Sweden) and cloned in pRT100mod/eYFP vector via restriction digestion (BamHI and NcoI) and subsequent ligation (Baudisch et al. 2014). The above gene constructions, comprising promoter and terminator regions (*CaMV35S::Gene of interest:GFP1-10/GFP11$_{x7}$::t35S*), were sub-cloned into a *Golden Gate* compatible pLSU4GG binary vector with 'standard' BsaI restriction/ligation *Golden Gate* reaction (see below **g**).

(d) Construction of the PlaMiNGo toolkit (Chapter 3)

The modular cloning principle and DNA fragments of the Plant Parts I and II toolkits were used for vector construction (Weber et al., 2011; Engler et al., 2014; Gantner et al., 2018). The modules utilized for cloning of *Golden Gate*-based vectors are illustrated in Figure 3.6. In this case *Golden Gate* reactions were performed with 200 ng of each DNA module with the following conditions: 2.5 units of T4 DNA ligase (Thermo Scientific), 5 units of BsaI or BpiI (New England Biolabs), 30 cycles of incubation at 37°C for 2 min and 16°C for 5 min, final denaturation at 80°C for 10 min. When required, the restriction/ligation reactions were subsequently supplemented with fresh ligation buffer and ligase for terminal ligation, and incubated for ≥ 3h at 16°C. Ligation mixtures were transformed into Dh10b or ccdB survival II cells (Thermo Scientific) and grown on LB plates with appropriate selective medium. PCR primers for Level 0 modules and gene sequences are summarized in Table 6.5 and Appendix.

(d) Cloning of candidate proteins into PlaMiNGo vectors (Chapter 3)

The candidate targeting signals (TyrRS$_{1-91}$, FNR$_{1-55}$, mtRi$_{1-100}$, ATPS$_{1-100}$, GCS$_{1-100}$, GAPDH$_{1-100}$, Gtred$_{1-100}$, RbcS$_{1-79}$ and CoxIV$_{1-29}$) were amplified from the corresponding cDNA templates (Nelson et al., 2007; Berglund et al., 2009; Baudisch et al., 2014) and cloned via 'standard' *Golden Gate* reaction (section 5.2.1g) into PLaMiNGo vectors in exchange for a ccdB negative selection cassette. Overhangs of the fragment to be cloned were A|ATG at the 5´ end and T|TCG at the 3´ end. Two additional nucleotides were inserted in some of the fusions to maintain the reading frame resulting in an additional codon for an alanine residue.

(e) Cloning for complementation of the immutans phenotype (Chapter 3)

The gene sequence carrying full length (1-295 AA; 885 bp) or mature PTOX (57-295 AA; 717 bp) was amplified from a cDNA clone provided by Steven Rodermel (Iowa State University, USA). The PTOX full-length gene and mature PTOX (lacking the transit peptide) were cloned into pRT100mod vectors downstream to CaMV35 promoter via restriction digestion (NcoI and

BsrGI) and subsequent ligation. The mature PTOX was cloned via RF cloning reaction into pRT100mod-based vectors downstream of the gene sequences coding for either GCS_{1-100}, $GrpE_{1-100}$, $ATPS_{1-100}$ or $Ef\text{-}Tu_{1-100}$. The candidate gene constructions with promoter and terminator (*CaMV35S::Candidate:PTOXmat::t35S*) were then cloned into a *Golden Gate* compatible pLSU4GG binary vector via standard *Golden Gate* cloning reaction (Appendix II).

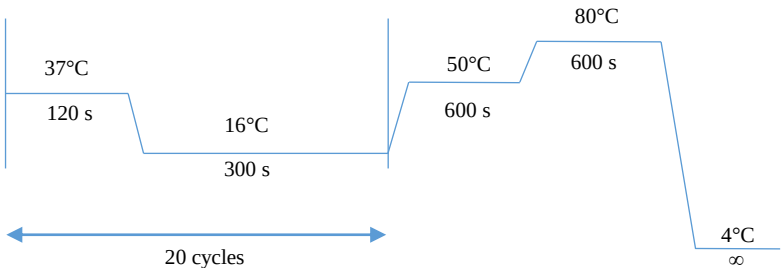
(f) Construction of Crispr/Cas9 vector and generation of Arabidopsis thaliana tatA mutants (Chapter 4)

Multiplex genome editing vector and protocol was used as described in Ordon et al. (2016). Two small guide RNAs (sgRNA) complementary to 20 nucleotides in exon I and exon II region of *Arabidopsis thaliana* TatA gene were designed using CHOPCHOP online tool (Kornel et al., 2016) (http://chopchop.cbu.uib.no/) and no off-targets were detected when analyzed by REGEN online tools (Bae et al., 2014) (http://www.rgenome.net/cas-offinder/). The oligonucleotide pairs were custom synthesized in a way that they expose 5' and 3' overhangs after hybridization. The oligonucleotides were hybridized by 5 min melting and annealing at 80°C and 16°C respectively. The hybridized oligos were inserted in vectors pDGE5 and pDGE8 via *Golden Gate* reaction (Ordon et al., 2016) constructing two new vectors namely pDGE5:oligo3TatA and pDGE8:oligo16TatA. Both oligo constructions were then sub-cloned in a binary vector pDGE4 via 'standard' *Golden Gate* reaction The final vector, pDGE4:TatACrispr is comprised of gene sequences coding for Cas9 driven by PcU6 promoter and a scaffold guide RNA expressed under control of AtUbi10 promoter. This binary vector pDGE4:TatACrispr was used to transform Arabidopsis thaliana ecotype Columbia-0 plants using floral dip transformation method as described by Davis et al. (2009). The genotyping of plants was performed with *TatA* screening primers and *hCas9* specific primers (see section 4.2.5 for details).

(g) Standard Golden Gate *Reaction*

Components	Amount
Vector	~50 ng
PCR Product	~50-100 ng
Ligase Buffer	1.5 µl
10× BSA (Invitrogen)	1.5 µl
BsaI/BpiI Enzyme	0.75 µl
T4 DNA Ligase	0.75 µl
ddH2O	Up to 15 µl

Reaction Cycle



1.5 µl of reaction mix was used to transform *E. coli* DH5α electro-competent cells.

(h) Vectors newly generated in this study

Table 5.7 List of vectors generated in this study.

Name of Vector	Parent Vector	Insert	Selection (Plant)	Cloning Method*
Comparative Analysis (Chapter 2)				
pRT100mod:TyrRSeYFP	pRT100mod	TyrRS1-91	Amp	RD (BamHI/NcoI)
pLSU4:TyrRSeYFP	pLSU4GG	Prom:TyrRS91:eYFP:Term	Kan (Hyg)	GG (BsaI)
pRT100mod:TyrRmCherry	pRT100mod: TyrRSeYFP	mCherry	Amp	RF (eYFP replacement)
pLSU4:TyrRmCherry	pLSU4GG	Prom:TyrRS91:mCherry:Term	Kan (Hyg)	GG (BsaI)
pCB302:EftueYFP	pCB302	Ef-Tu1-100	Kan (BASTA)	RD (PstI)
pCB302:GCSeYFP	pCB302	GCS1-100	Kan (BASTA)	RD (PstI)
pCB302:GrpEeYFP	pCB302	GrpE1-100	Kan (BASTA)	RD (PstI)
pLSU4:PDFeYFP	pLSU4GG	PDF1-100	Kan (Hyg)	GG (BsaI)
self-assembling split-GFP (Chapter 3)				
pRT100mod:mtrGFP10	pRT100mod: mtReYFP	GFP1-10	Amp	RD (BsaII/XbaI) (eYFP replacement)
pRT100mod:mtrGFP11x7	pRT100mod: mtReYFP	GFP11x7	Amp	RD (NcoI/XbaI) (eYFP replacement)
pRT100mod:FNRGFP10	pRT100mod: FNReYFP	GFP1-10	Amp	RD (BsaII/XbaI) (eYFP replacement)
pRT100mod:FNRGFP11x7	pRT100mod: FNReYFP	GFP11x7	Amp	RD (NcoI/XbaI) (eYFP replacement)
pLSU4:mtrGFP10	pLSU4GG	Prom:mtr100:GFP1-10:Term	Kan (Hyg)	GG (BsaI)
pLSU4:mtrGFP11x7	pLSU4GG	Prom:mtr100:GFP11x7:Term	Kan (Hyg)	GG (BsaI)
pLSU4:FNRGFP10	pLSU4GG	Prom:FNR55:GFP1-10:Term	Kan (Hyg)	GG (BsaI)
pLSU4:FNRGFP11x7	pLSU4GG	Prom:FNR55:GFP11x7:Term	Kan (Hyg)	GG (BsaI)
pRT100mod:TyrGFP11x7	pRT100mod: TyrRSeYFP	GFP11x7	Amp	RD (NcoI/XbaI) (eYFP replacement)

Name of Vector	Parent Vector	Insert	Selection (Plant)	Cloning Method*
pLSU4:TyrGFP11x7	pLSU4GG	Prom:Tyr91:GFP11 x7:Term	Kan (Hyg)	GG (BsaI)
pRT100mod:TyrGFP11x3	pRT100mod: TyrRSeYFP	GFP11x3	Amp	RD (NcoI/XbaI) (eYFP replacement)
pLSU4:TyrGFP11x3	pLSU4GG	Prom:Tyr91:GFP11 x3:Term	Kan (Hyg)	GG (BsaI)
pRT100mod:TyrGFP11x1	pRT100mod: TyrRSeYFP	GFP11x1	Amp	RD (NcoI/XbaI) (eYFP replacement)
pLSU4:TyrGFP11x1	pLSU4GG	Prom:Tyr91:GFP11 x1:Term	Kan (Hyg)	GG (BsaI)
pRT100mod:GrpGFP11x7	pRT100mod: GrpEeYFP	GFP11x7	Amp	RF (eYFP replacement)
pLSU4:GrpGFP11x7	pLSU4GG	Prom:GrpE100:GFP 11x7:Term	Kan (Hyg)	GG (BsaI)
pRT100mod:PDFGFP11x7	pRT100mod: PDFeYFP	GFP11x7	Amp	RD (NcoI/XbaI)
pLSU4:PDFGFP11x7	pLSU4GG	Prom:PDF100:GFP 11x7:Term	Kan (Hyg)	GG (BsaI)
pRT100mod:pTatA100GF P11x7	pRT100mod: pTatA100eYFP	GFP11x7	Amp	RF (eYFP replacement)
pLSU4:pTatA100GFP11x7	pLSU4GG	Prom:pTatA100:GF P11x7:Term	Kan (Hyg)	GG (BsaI)
pLaMiNGo Toolkit				
pTEI119	pAGM1276	FNR55:GFP1-10	Spec	GG (BpiI)
pTEI120	pAGM1276	mtRi100: GFP1-10	Spec	GG (BpiI)
pTEI121	pJOG387	GFP1-10	Spec	GG (BpiI)
pTEI122	pJOG267	ccdB	Spec	GG (BpiI)
pTEI123	pAGM1301	GFP11x3	Spec	GG (BpiI)
pTEI124	pAGM1301	GFP11x7	Spec	GG (BpiI)
pTEI126	pICH47742	Prom:FNR55: GFP1-10:Term	Carb	GG (BsaI)
pTEI127	pICH47742	Prom:mtRi100: GFP1-10:Term	Carb	GG (BsaI)
pTEI128	pICH47751	Prom:ccdB: GFP11x3:Term	Carb	GG (BsaI)
pTEI129	pICH47751	Prom:ccdB: GFP11x7:Term	Carb	GG (BsaI)
pTEI161	pICH47751	Prom:ccdB: GFP11x1:Term	Carb	GG (BsaI)
pTEI176	pICH47742	Prom:ccdB (AATG|TTCG): GFP1-10:Term	Carb	GG (BsaI)
pTEI177	pICH47742	Prom:ccdB (CCAT|AATG): GFP1-10:Term	Carb	GG (BsaI)
pLaNGo11-1	pAGM4723	Prom:FNR55: GFP1-10: Term::Prom:ccdB: GFP11x1:Term	Kan	GG (BpiI)
pLaNGo11-3	pAGM4723	Prom:FNR55: GFP1-10: Term::Prom:ccdB: GFP11x3:Term	Kan	GG (BpiI)

Name of Vector	Parent Vector	Insert	Selection (Plant)	Cloning Method*
pLaNGo11-7	pAGM4723	Prom:FNR55: GFP1-10: Term::Prom:ccdB: GFP11x7:Term	Kan	GG (BpiI)
pMiNGo11-1	pAGM4723	Prom:mtRi100: GFP1-10: Term::Prom:ccdB:G FP11x1:Term	Kan	GG (BpiI)
pMiNGo11-1	pAGM4723	Prom:mtRi100: GFP1-10: Term::Prom:ccdB:G FP11x1:Term	Kan	GG (BpiI)
pMiNGo11-3	pAGM4723	Prom:mtRi100: GFP1-10: Term::Prom:ccdB:G FP11x3:Term	Kan	GG (BpiI)
pMiNGo11-7	pAGM4723	Prom:mtRi100: GFP1-10: Term::Prom:ccdB:G FP11x7:Term	Kan	GG (BpiI)
pLaNGo11-1:Tyrs	pLaNGo11-1	TyrRS1-91	Kan	GG (BsaI)
pLaNGo11-3:Tyrs	pLaNGO11-3	TyrRS1-91	Kan	GG (BsaI)
pLaNGo11-7:Tyrs	pLaNGO11-7	TyrRS1-91	Kan	GG (BsaI)
pMiNGo11-1:Tyrs	pMiNGo11-1	TyrRS1-91	Kan	GG (BsaI)
pMiNGo11-3Tyrs	pMiNGo11-3	TyrRS1-91	Kan	GG (BsaI)
pMiNGo11-7Tyrs	pMiNGo11-7	TyrRS1-91	Kan	GG (BsaI)
pLaNGo11-7:FNR	pLaNGO11-7	FNR1-55	Kan	GG (BsaI)
pMiNGo11-1:FNR	pMiNGo11-1	FNR1-55	Kan	GG (BsaI)
pLaNGo11-7:mtRi	pLaNGO11-7	mtRi1-100	Kan	GG (BsaI)
pMiNGo11-1:mtRi	pMiNGo11-1	mtRi1-100	Kan	GG (BsaI)
pLaNGo11-7:GCS	pLaNGO11-7	GCS1-100	Kan	GG (BsaI)
pMiNGo11-1:GCS	pMiNGo11-1	GCS1-100	Kan	GG (BsaI)
pLaNGo11-7:Gtred	pLaNGO11-7	Gtred1-100	Kan	GG (BsaI)
pMiNGo11-1:Gtred	pMiNGo11-1	Gtred1-100	Kan	GG (BsaI)
pLaNGo11-7:GAPDH	pLaNGO11-7	GAPDH1-100	Kan	GG (BsaI)
pMiNGo11-1:GAPDH	pMiNGo11-1	GAPDH1-100	Kan	GG (BsaI)
pLaNGo11-7:ATPS	pLaNGO11-7	ATPS1-100	Kan	GG (BsaI)
pMiNGo11-7:ATPS	pMiNGo11-7	ATPS1-100	Kan	GG (BsaI)
pLaNGo11-7:CoxIV	pLaNGO11-7	CoxIV1-29	Kan	GG (BsaI)
pMiNGo11-7:CoxIV	pMiNGo11-7	CoxIV1-29	Kan	GG (BsaI)
pLaNGo11-7:SSU1-79	pLaNGO11-7	SSU1-79	Kan	GG (BsaI)
pMiNGo11-7:SSU1-79	pMiNGo11-7	SSU1-79	Kan	GG (BsaI)
Phenotype complementation immutans *(Chapter 3)*				
pRT100mod:PTOX	pRT100mod	PTOX	Amp	RD (NcoI/BsrGI)
pLSU4:PTOX	pLSU4GG	Prom:PTOX:Term	Kan (Hyg)	GG (BsaI)
pRT100PTOXmat	pRT100mod	PTOX57-351	Amp	RD (NcoI/BsrGI)
pLSU4:PTOXmat	pLSU4GG	Prom:PTOXmat:Term	Kan (Hyg)	GG (BsaI)

Name of Vector	Parent Vector	Insert	Selection (Plant)	Cloning Method*
pRT100mod:GSCPtxmat	pRT100mod: GCSeYFP	PTOX57-351	Amp	RD (NcoI/BsrGI) (eYFP replacement)
pLSU:GCSPtxmt	pLSU4GG	Prom:GCS100:PTOXmat:Term	Kan (Hyg)	GG (BsaI)
pRT100:ATPSPtxmt	pRT100mod: ATPSeYFP	PTOX57-351	Amp	RD (NcoI/BsrGI) (eYFP replacement)
pLSU4:ATPSPtxmt	pLSU4GG	Prom:ATPS100:PTOXmat:Term	Kan (Hyg)	GG (BsaI)
pRT100mod:EfTuPtxmt	pRT100mod: GSCPtxmat	Ef-Tu1-100	Amp	RF (GCS replacement)
pLSU4:EfTuPtxmt	pLSU4GG	Prom:Ef-Tu100:PTOXmat:Term	Kan (Hyg)	GG (BsaI)
pRT100mod:GrpPtxmt	pRT100mod: GSCPtxmat	GrpE1-100	Amp	RF (GCS replacement)
pLSU4:GrpPtxmt	pLSU4GG	Prom:GrpE100:PTOXmat:Term	Kan (Hyg)	GG (BsaI)
Targeting specificity of TatA (Chapter 4)				
pLSU4:AtTatA100eYFP	pLSU4GG	Prom:AtTatA100:eYFP:Term	Kan (Hyg)	GG (BsaI)
pRT100:AtTatAfulleYFP	pRT100mod: eYFP	AtTatA	Amp	RD (BamHI/NcoI)
pLSU4:AtTatAfulleYFP	pLSU4GG	Prom:AtTatA:eYFP:Term	Kan (Hyg)	GG (BsaI)
pDGE5:oligo3TatA	pDGE5	CrisprTatA Oligo3	Amp	GG (BpiI)
pDGE8:oligo16TatA	pDGE8	CrisprTatAOligo16	Amp	GG (BpiI)
pDGE4:AtTaA	pDGE4	Oligo3:Oligo16	Spect	GG (BsaI)

*GG- Golden Gate Cloning, RD- Restriction digestion and ligation, RF- restriction free cloning (Bond and Naus, 2012); Amp- Ampicillin, Kan- Kanamycin, Hyg- Hygromycin, Spect- Spctinomycin; Vector construction of PlaMiNGo toolkit is presented in Figure 3.6

6.2.1 Cultivation of Plants

(a) Arabidopsis thaliana (Col-0 background)

The seeds from respective *Arabidopsis thaliana* plant lines were stratified al least for 3 days in the cold room (temp. 4-5° C) before germination. The seeds were germinated on soil (Substrat I, Klasmann-Deilmann GmbH, Germany) mixed with vermiculite in a 3:1 ratio and cultivated for 3-4 weeks in a plant cultivation room with 8/16 h light-dark cycle, 20-22° C temperature and approximately 150 µmol m^{-2} s^{-1} light. For induction of flowering, the plants were transferred in the green house with long day, 16/8 h light dark cycle at 24° C.

(b) Arabidopsis thaliana (Immutans mutants)

The seeds from *Arabidopsis thaliana immutans* mutant lines were stratified at least for 3 days in the cold room (temp. 4-5° C) before germination. The seeds were germinated on soil (Substrat I,

Klasmann-Deilmann GmbH, Germany) with vermiculite in a 3:1 ratio and cultivated for 3 weeks in a plant cultivation room with 8/16 h light-dark cycle, 20-22° C temperature and approximately 5 µmol m^{-2} s^{-1} light. Later, the plantlets were transferred to a chamber with approximately 50 µmol m^{-2} s^{-1} light and cultivated for 7-8 weeks. Afterwards, the plants were transferred to green house with 16/8 h light and dark cycle, 24° C temperature and >150 µmol m^{-2} s^{-1} light for induction of flowering.

(c) Nicotiana benthamiana

The *Nicotiana benthamiana* seeds were germinated on soil (Substrat I, Klasmann-Deilmann GmbH, Germany) and cultivated in the green house with 16/8 h light dark cycle and 23-25° C temperature.

(d) Pisum sativum

The *Pisum sativum* seeds were germinated on soil (Profi Substrat, Einheitserde Werkverband e.V., Germany) and cultivated in the green house with 16/8 h light dark cycle and 20-22° C temperature.

6.2.2 Transient transformation

(a) Particle Bombardment

The particle bombardment of 3-5 week old leaves from Arabidopsis thaliana cv. Col-0 was performed as described in Rödiger et al. (2011). For a single shot, 0.2 mg gold particles of 0.6 µm diameter were used as a DNA carrier and approx. 300 ng of plasmid DNA were precipitated onto gold particle and bombarded on adaxial side of 4-5-week-old *Arabidopsis thaliana* leaves. For localization studies in *Pisum sativum* (pea), 1-week-old leaves were utilized. After bombardment, the leaves were incubated overnight in dark for approx. 16-20 h and later analyzed by confocal laser scanning microscope.

(b) Protoplast Transformation

Protoplasts from *Arabidopsis thaliana* leaves were isolated following the "Tape *Arabidopsis* sandwich" method described by Wu et al. (2009) (Appendix II). For transformation, the isolated protoplasts were incubated for 5 min with 10 µg of DNA construct, 0.2 M Mannitol, 0.1 M $CaCl_2$, and 40% PEG-4000, washed twice with W5 buffer (154 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, 5 mM Glucose and 2 mM MES, pH 5.7), resuspended in modified Mg solution (0.4 M mannitol, 15 mM MgCl2, and 4 mM MES, pH 5.7) and further incubated for 16-20 h under continuous 150 µmol m^{-2} s^{-1} light. The transformed protoplasts were later placed on 30 mm Glass

bottom (No. 1.5) Microwell petridishes (MatTek Corporation, Ashland USA) and imaged with CLSM.

(c) Agrobacterium mediated transient transformation of Nicotiana benthamiana

Agrobacterium strains carrying the candidate gene constructs were harvested after incubation for 72 h at 28°C from LB agar plates supplemented with the appropriate antibiotic. After suspension in infiltration medium (10 mM MgCl2, 10 mM MES, 150 µM acetosyringone) the cultures were adjusted to OD600 = 0.8, incubated for 3 h at room temperature and infiltrated with a needleless syringe into the lower epidermis of fully expanded leaves from 6 to 8 weeks old *Nicotiana benthamiana* plants. For co-infiltration experiment, each bacterial strain was adjusted to OD_{600} = 0.8 and mixed in a 1:1 ratio prior to infiltration. After incubation for 3 days with a 16/8 h light-dark cycle, protein localization was analyzed using confocal laser scanning microscopy (CLSM). For staining epidermal cell mitochondria, 0.1 µM of MitoTracker Orange Solution (0.1 µM MitoTracker in infiltration medium) was infiltrated in lower epidermis of leaf tissue with needleless syringe infiltration, 15 min before imaging.

6.2.3 Generation of *Arabidopsis* thaliana stable transgenic lines

Wild-type Col-0 ecotype of *Arabidopsis thaliana*, *Arabidopsis thaliana tatA(+/-)* mutant and *Arabidopsis thaliana immutans* mutant plants were transformed using the floral dip transformation method described by Davis et al. (2009). Transformed plants were selected by either spraying with 0.1% BASTA (for pCB302 based constructs) or on ½ MS plates containing 30 µg/ml Hygromycin-B (for pLSU4GG-based constructs). For each construct, at least three independent T1 and T2 transgenic lines were analysed.

6.2.5 DNA Extraction

The plasmid DNA from *E. coli* or *Agrobacterium* was extracted using **NucleoSpin® Plasmid** kit (Macherey Nagel). For extraction of genomic DNA from plant leaf tissue, cotyledons or leaf disks were crushed in 400 µl of DNA extraction buffer (200 mM Tris-HCl, pH 7.5, 250 mM NaCl, 25mM EDTA, 0.5% SDS), the DNA was precipitated from supernatant after addition of equal volume isopropanol and incubation at -20° C for at least 30 minutes. The precipitated DNA was washed with 70% ethanol and dissolved in 50 µl of sterile water. For genotyping, 1.0 µl of genomic DNA was used as a PCR template to amplifying the specific fragments.

6.2.4 Confocal laser scanning microscopy (CLSM)

Confocal laser scanning microscopy was carried out with a Zeiss LSM780 Confocal Imaging System. For the emission of fluorescence signals specimens were excited with either 488 nm (eGFP and *sa*split-GFP), 514 nm (eYFP), 561 (MitoTracker Orange and mCherry), or 633 nm

(chlorophyll) laser usually with, unless otherwise mentioned, 2% of full laser power. Images were collected using filters ranging from 493 to 598 nm (eGFP), 519–620 nm (eYFP), 574-617 (MitoTracker Orange), 607-630 (mCherry) or 647–721 nm (chlorophyll). For GFP/YFP dual channel imaging, the 493-518 (GFP) and 519-620 (YFP) filter ranges were used. Image acquisition was done in several Z-stacks with either 20X or 40X objectives. In most instances, either (x: 1024, y: 1024, z: 5, 16-bit) or (x: 512, y: 512, z: 10, 16-bit) frame dimensions were used for image acquisition with 20X and 40X objectives respectively. The Pinhole was set to one airy units (A.U.) and essentially identical settings were used for acquisition of all images. Unless otherwise mentioned, the images were presented as maximum intensity projection.

6.2.5 Image processing and signal quantification

Brightness and contrast of the images were later adjusted in order to better visualize the fluorescence signals. In case where a quantitative data is presented, the brightness is equally adjusted for all images. All images were processed using ZEN software (Carl Zeiss, Jena) and Inkscape™ (GPL, v3). Quantification of the signals was performed with raw images using the Fiji program (Schindelin et al., 2012). For quantification of signal intensities to compare the *sa*split-GFP vectors (see section 3.2.2), infiltration of all relevant constructs was carried out on different spots of the same leaf. At least three images from each infiltration spot were used for quantification. The image acquisition was done with the 20X objective in 7 to 8 Z-stacks covering the epidermal cell layer and later stacked to project the maximum intensities. The mean grey values of stacked images were calculated using the 'Measure' option of Fiji and further utilized for comparison of the fluorescence signal strengths in arbitrary units (A.U.).

6.2.4 Western blot (Immunoblot)

For detection of PTOX protein, the western blot was performed with total protein extract from 10 mg of leaf tissue isolated with SDS containing protein isolation buffer (125 mM Tris-HCl, pH8.8, 1% w/v SDS, 10 v/v glycerol and 2 mM DTT). The crude protein extract was separated by SDS-PAGE on 12.5% polyacrylamide gel (8.3 x 6.4 cm) and transferred onto PVDF membranes (pore size- 0.2 µm) via semidry blotting (45 min. at 1.5 mA/cm^2) using cathode and anode running buffer. Prehybridization and hybridization with primary (1:4000) and secondary antibody (1:30,000) was carried out with modified blocking solution 1 (1x PBS, 5% Milk powder, 3% BSA and 0.1% Tween-20). After hybridisation, the membrane was washed with blocking solution 2 (1x PBS containing 0.1% v/v Tween 20). The signals were developed with standard ECL (enhanced chemiluminescence) solution (1 M Tris-HCl, pH 8.5, 250 mM Luminol, 90 mM Coumaric acid, 30% H_2O_2) and exposed to X-ray film for 2 min.

References

Abel S, Theologis A (1994) Transient transformation of Arabidopsis leaf protoplasts: a versatile experimental system to study gene expression. Plant Journal. 5:421-427.

Adams DG (2000) Symbiotic interactions. In The ecology of cyanobacteria (Pages 523-561). Springer. Dordrecht.

Agarraberes FA, Dice JF (2001) Protein translocation across membranes. Biochimica et Biophysica Acta (BBA)-Biomembranes.1513:1-24.

Aluru MR Bae H Wu D and Rodermel SR (2001) The Arabidopsis immutans mutation affects plastid differentiation and the morphogenesis of white and green sectors in variegated plants. Plant Physiology. 127:67-77.

Andresen M Schmitz-Salue R and Jakobs S (2004) Short tetracysteine tags to β-tubulin demonstrate the significance of small labels for live cell imaging. Molecular Biology of the Cell. 15:5616-5622.

Bandlow W, Strobel G, Zoglowek C, Oechsner U, Magdolen V (1988) Yeast adenylate kinase is active simultaneously in mitochondria and cytoplasm and is required for non-fermentative growth. FEBS Journal. 178:451-457.

Barnett JP, Robinson C, Scanlan DJ, & Blindauer CA (2011) The Tat protein export pathway and its role in cyanobacterial metalloprotein biosynthesis. FEMS Microbiology Letters. 325:1-9.

Barton KA Schattat MH, Jakob T, Hause G, Wilhelm C, Mckenna JF, Máth C, Runions J Van Damme D and Mathur J (2016) Epidermal pavement cells of Arabidopsis have chloroplasts. Plant Physiology. 171:723-726.

Baudisch B, Klösgen RB (2012) Dual targeting of a processing peptidase into both endosymbiotic organelles mediated by a transport signal of unusual architecture. Molecular Plant. 5:494-503.

Baudisch B, Langner U, Garz I, Klösgen RB (2014) The exception proves the rule? Dual targeting of nuclear-encoded proteins into endosymbiotic organelles. New Phytologist. 201:80–90.

Bauwe H, & Kolukisaoglu Ü (2003) Genetic manipulation of glycine decarboxylation. Journal of Experimental Botany. 54:1523-1535.

Berglund AK, Pujol C, Duchene AM and Glaser E (2009) Defining the determinants for dual targeting of amino acyl-tRNA synthetases to mitochondria and chloroplasts. Journal of Molecular Biology. 393:803-814.

Birch RG (1997) Plant Transformation: Problems and Strategies for Practical Application. Annual Review Plant Physiology and Plant Molecular Biology. 48:297–326.

Bock R and Timmis JN (2008) Reconstructing evolution: gene transfer from plastids to the nucleus. Bioessays. 30:556-566.

Bolliger L, Deloche O, Glick BS, Georgopoulos C, Jenö P, Kronidou N, Horst, M, Morishima N and Schatz G (1994) A mitochondrial homolog of bacterial GrpE interacts with mitochondrial hsp70 and is essential for viability. The EMBO Journal. 13:1998-2006.

Bölter B (2018). En route into chloroplasts: preproteins' way home. Photosynthesis Research. 138: 263–275.

Bond SR and Naus CC (2012) RF-Cloning.org: an online tool for the design of restriction-free cloning projects. Nucleic Acids Research. 40:W209-W213.

Braun HP, Schmitz UK (1999) The protein-import apparatus of plant mitochondria. Planta. 209:267-274.

Bruce BD (2000) Chloroplast transit peptides: structure, function and evolution. Trends in Cell Biology. 10:440-447.

Cabantous S, Terwilliger TC, Waldo GS (2005) Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. Nature Biotechnology. 23:102–107.

Cabantous S, Waldo GS (2006) In vivo and in vitro protein solubility assays using split GFP. Nature Methods. 3:845.

Cabantous S, Nguyen H, Pedelacq J, Koraïchi F, Chaudhary A, Ganguly K, Lockard MA, Favre G, Terwilliger TC and Waldo GS (2013) A new protein-protein interaction sensor based on tripartite split-GFP association. Scientific Reports. 3:2854.

Carol P, Stevenson D, Bisanz C, Breitenbach J, Sandmann G, Mache R, Coupland G and Kuntz M (1999) Mutations in the Arabidopsis gene IMMUTANS cause a variegated phenotype by inactivating a chloroplast terminal oxidase associated with phytoene desaturation. The Plant Cell. 11:57-68.

Carrie C, Murcha MW, Kuehn K, Duncan O, Barthet M, Smith PM, Eubel H, Meyer E, Day DA, Millar AH, Whelan J (2008) Type II NAD(P)H dehydrogenases are targeted to mitochondria and chloroplasts or peroxisomes in Arabidopsis thaliana. FEBS Letters. 582:3073–3079.

Carrie C, Kühn K, Murcha MW, Duncan O, Small ID, O'Toole N and Whelan J (2009) Approaches to defining dual-targeted proteins in Arabidopsis. The Plant Journal. 57:1128-1139.

Carrie C, Small I (2013) A reevaluation of dual-targeting of proteins to mitochondria and chloroplasts. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research.1833:253-259.

Carrie C, Whelan J (2013) Widespread dual targeting of proteins in land plants: When, where, how and why. Plant Signaling and Behavior. 8:e25034.

Carrie C, Weißenberger S, & Soll J (2016) Plant mitochondria contain the protein translocase subunits TatB and TatC. Journal of Cell Science. 129:3935-3947.

Chabregas SM, Luche DD, Farias LP, Ribeiro AF, Van Sluys MA, Menck CF, & Silva-Filho MC (2001) Dual targeting properties of the N-terminal signal sequence of Arabidopsis thaliana THI1 protein to mitochondria and chloroplasts. Plant Molecular Biology. 46:639–650.

Chaddock AM, Mant A, Karnauchov I, Brink S, Herrmann RG, Klösgen RB, & Robinson C (1995) A new type of signal peptide: central role of a twin-arginine motif in transfer signals for the delta pH-dependent thylakoidal protein translocase. The EMBO Journal. 14:2715-2722.

Chen YL, Chen LJ, Chu CC, Huang PK, Wen JR, & Li HM (2018) TIC236 links the outer and inner membrane translocons of the chloroplast. Nature. 564:125.

Cheng NH (2008) AtGRX4, an Arabidopsis chloroplastic monothiol glutaredoxin, is able to suppress yeast grx5 mutant phenotypes and respond to oxidative stress. FEBS Letters. 582:848-854.

Chiu WL, Niwa Y, Zeng W, Hirano T, Kobayashi H and Sheen J (1996) Engineered GFP as a vital reporter in plants. Current Biology. 6:325-330.

Chow KS, Singh DP, Roper JM and Smith AG (1997) A single precursor protein for Ferrochelatase-I from Arabidopsis is imported in vitro into both chloroplasts and mitochondria. Journal of Biological Chemistry. 272:27565-27571.

Christensen AC, Lyznik A, Mohammed S, Elowsky CG, Elo A, Yule R, Mackenzie SA (2005) Dual-domain, dual-targeting organellar protein presequences in Arabidopsis can use non-AUG start codons. Plant Cell. 17:2805–2816

Cline K, Ettinger WF & Theg SM (1992) Protein-specific energy requirements for protein transport across or into thylakoid membranes. Two lumenal proteins are transported in the absence of ATP. Journal of Biological Chemistry. 267:2688-2696.

Cline K, Mori H (2001). Thylakoid ΔpH-dependent precursor proteins bind to a cpTatC–Hcf106 complex before Tha4-dependent transport. Journal of Cell Biology. 154:719-730.

Coruzzi G, Broglie R, Edwards C and Chua NH (1984). Tissue-specific and light-regulated expression of a pea nuclear gene encoding the small subunit of ribulose-1, 5-bisphosphate carboxylase. The EMBO Journal. 3:1671-1679.

Creissen G, Reynolds H, Xue Y and Mullineaux P (1995) Simultaneous targeting of pea glutathione reductase and of a bacterial fusion protein to chloroplasts and mitochondria in transgenic tobacco. The Plant Journal. 8:167-175.

Davis AM, Hall A, Millar, AJ, Darrah C and Davis SJ (2009) Protocol: Streamlined sub-protocols for floral-dip transformation and selection of transformants in Arabidopsis thaliana. Plant Methods. 5:3.

De Greve H., Dhaese P, Seurinck J, Lemmers M, Van Montagu M, Schell J (1982) Nucleotide sequence and transcript map of the Agrobacterium tumefaciens Ti plasmid-encoded octopine synthase gene. Journal of Molecular and Applied Genetics. 1:499–511.

Dinkins, R. D., Conn, H. M., Dirk, L. M., Williams, M. A., & Houtz, R. L. (2003). The Arabidopsis thaliana peptide deformylase 1 protein is localized to both mitochondria and chloroplasts. Plant Science. 165:751-758.

Donzeau M, Kaldi K, Adam A, Paschen S, Wanner G, Guiard B (2000) Tim23 links the inner and outer mitochondrial membranes. Cell. 101:401–412.

Douce R, Bourguignon J, Neuburger M, and Rébeillé F (2001) The glycine decarboxylase system: a fascinating complex. Trends in Plant Science. 6:167-176.

Duchêne A-M, Giritch A, Hoffmann B, Cognat V, Lancelin D, Peeters NM, Zaepfel M, Maréchal-Drouard L, Small ID (2005) Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in Arabidopsis thaliana. Proceedings of National Academy of Sciences. 102:16484-16489.

Dupree P, Pwee KH, Gray JC (1991) Expression of photosynthesis gene-promoter fusions in leaf epidermal cells of transgenic tobacco plants. Plant Journal. 1:115–120.

Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting Subcellular Localization of Proteins Based on their N-terminal Amino Acid Sequence. Journal of Molecular Biology. 300:1005–1016.

Emmermann M, Clericus M, Braun HP, Mozo T, Heins L, Kruft V and Schmitz UK (1994) Molecular features, processing and import of the Rieske iron-sulfur protein from potato mitochondria. Plant Molecular Biology. 25:271-281.

Engler C, Youles M, Gruetzner R, Ehnert TM, Werner S, Jones JD, Patron, NJ and Marillonnet S (2014) A golden gate modular cloning toolbox for plants. ACS Synthetic Biology. 3:839-843.

Erickson JL, ZieglerJ, Guevara D, Abel S, Klösgen RB, Mathur J, Rothstein SJ and Schattat MH (2014) Agrobacterium-derived cytokinin influences plastid morphology and starch accumulation in Nicotiana benthamiana during transient assays. BMC Plant Biology. 14:127.

Erickson, JL, Kantek M, and Schattat MH (2017) Plastid-nucleus distance alters the behavior of stromules. Frontiers in Plant Science. 8:1135.

Faraco M, Di Sansebastiano GP, Spelt K, Koes RE and Quattrocchio FM (2011) One protoplast is not the other!. Plant Physiology. 156:474-478.

Flores-Pérez Ú, Jarvis P (2013) Molecular chaperone involvement in chloroplast protein import. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research. 1833:332-340.

Foudree A, Putarjunan A, Kambakam S, Nolan T, Fussell J, Pogorelko G and Rodermel S (2012) The mechanism of variegation in immutans provides insight into chloroplast biogenesis. Frontiers in Plant Science. 3:260.

Fu A, Aluru M, & Rodermel SR (2009) Conserved active site sequences in arabidopsis plastid terminal oxidase (PTOX) in vitro and in planta mutagenesis studies. Journal of Biological Chemistry. 284:22625-22632.

Fu J, Momčilović I, & Prasad PV (2012) Roles of protein synthesis elongation factor EF-Tu in heat tolerance in plants. Journal of Botany. 2012:835836.

Fulgosi H, & Soll J (2002) The chloroplast protein import receptors Toc34 and Toc159 are phosphorylated by distinct protein kinases. Journal of Biological Chemistry. 277:8934-8940.

Fuss J, Liegmann O, Krause K, Rensing SA (2013) Green targeting predictor and ambiguous targeting predictor 2: The pitfalls of plant protein targeting prediction and of transient protein expression in heterologous systems. New Phytologist. 200:1022–1033

Gantner J, Ordon J, Ilse T, Kretschmer C, Gruetzner R, Löfke C, Dagdas Y, Bürstenbinder K, Marillonnet S, Stuttmann J (2018) Peripheral infrastructure vectors and an extended set of plant parts for the Modular Cloning system. PloS One. 13:0197185.

García-Mata R, Bebök Z, Sorscher EJ, Sztul ES (1999) Characterization and dynamics of aggresome formation by a cytosolic GFP-chimera. Journal of Cell Biology. 146:1239–1254.

Garg SG, Gould SB (2016) The Role of Charge in Protein Targeting Evolution. Trends in Cell Biology. 26:894–905.

Ge C, Spånning E, Glaser E, Wieslander Å (2014) Import determinants of organelle-specific and dual targeting peptides of mitochondria and chloroplasts in Arabidopsis thaliana. Molecular Plant. 7:121–136.

Gelvin SB (2003) Agrobacterium-mediated plant transformation: the biology behind the "gene-jockeying" tool. Microbiology and Molecular Biology Reviews. 67:16-37.

Ghifari AS, Gill-Hille M, & Murcha MW (2018) Plant mitochondrial protein import: The ins and outs. Biochemical Journal. 475: 2191-2208.

Giepmans BN, Adams SR, Ellisman MH and Tsien RY (2006) The fluorescent toolbox for assessing protein location and function. Science. 312:217-224.

Giglione C, Serero A, Pierre M, Boisson B and Meinnel T (2000) Identification of eukaryotic peptide deformylases reveals universality of N-terminal protein processing mechanisms. The EMBO Journal. 19:5916-5929.

Gray MW (1993) Origin and evolution of organelle genomes. Current Opinion in Genetics and Development. 3:884-890.

Grefen C and Blatt MR (2012) A 2in1 cloning system enables ratiometric bimolecular fluorescence complementation (rBiFC). Biotechniques. 53:311-314.

Guéra A, America T, van Waas M, & Weibeek PJ (1993) A strong protein unfolding activity is associated with the binding of precursor chloroplast proteins to chloroplast envelopes. Plant Molecular Biology. 23:309-324.

Guillaumot D, Lopez-Obando M, Baudry K, Avona A, Rigailla G, Falcon de Longeviallea A, Brochea B, Takenakad M, Berthoméa R, De Jaegere G, Delannoya E, and Lurina C (2017) Two interacting PPR proteins are major Arabidopsis editing factors in plastid and mitochondria. Proceedings of National Academy of Sciences. 114:8877-8882.

Gutensohn M, Fan E, Frielingsdorf S, Hanner P, Hou B, Hust B, & Klösgen RB (2006) Toc, Tic, Tat et al.: structure and function of protein transport machineries in chloroplasts. Journal of Plant Physiology. 163:333-347.

Hauer RS, Schlesier R, Heilmann K, Dittmar J, Jakob M, & Klösgen RB (2013) Enough is enough: TatA demand during Tat-dependent protein transport. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research. 1833:957-965.

Heazlewood JL, Tonti-Filippini JS, Gout AM, Day DA, Whelan J, & Millar AH (2004) Experimental analysis of the Arabidopsis mitochondrial proteome highlights signaling and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins. The Plant Cell. 16:241-256.

Hecker A, Wallmeroth N, Peter S, Michael RB, Harter K and Grefen C (2015) Binary 2in1 vectors improve in planta (co-) localisation and dynamic protein interaction studies. Plant Physiology. 168:776-87.

Hedtke B, Börner T, Weihe A (2000) One RNA polymerase serving two genomes. EMBO Reports. 1:435-440.

Hedtke B, Legen J, Weihe A, Herrmann RG, Börner T (2002). Six active phage-type RNA polymerase genes in Nicotiana tabacum. Plant Journal. 30:625-637.

Helm S, Dobritzsch D, Rödiger A, Agne B and Baginsky S (2014) Protein identification and quantification by data-independent acquisition and multi-parallel collision-induced dissociation mass spectrometry (MSE) in the chloroplast stroma proteome. Journal of Proteomics. 98:79-89.

Henry E, Toruño TY, Jauneau A, Deslandes L, Coaker GL (2017) Direct and indirect visualization of bacterial effector delivery into diverse plant cell types during infection. Plant Cell. 29:1555–1570

Highfield PE and Ellis RJ (1978) Synthesis and transport of the small subunit of chloroplast ribulose bisphosphate carboxylase. Nature. 271:420.

Hinnah SC, Hill K, Wagner R, Schlicher T, & Soll J (1997) Reconstitution of a chloroplast protein import channel. The EMBO Journal. 16:7351-7360.

Hinnah SC, Wagner R, Sveshnikova N, Harrer R, & Soll J (2002). The chloroplast protein import channel Toc75: pore properties and interaction with transit peptides. Biophysical Journal. 83:899-911.

Höglund A, Dönnes P, Blum T, Adolph HW, Kohlbacher O (2006) MultiLoc: Prediction of protein subcellular localization using N-terminal targeting sequences, sequence motifs and amino acid composition. Bioinformatics. 22:1158–1165.

Hooks KB, Turner JE, Graham IA, Runions J, Hooks MA (2012) GFP-tagging of Arabidopsis acyl-activating enzymes raises the issue of peroxisome-chloroplast import competition versus dual localization. Journal of Plant Physiology. 169:1631-1638.

Hooper CM, Castleden IR, Tanz SK, Aryamanesh N, Millar AH (2017) SUBA4: The interactive data analysis centre for Arabidopsis subcellular protein locations. Nucleic Acids Research. 45:D1064–D1074.

Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K (2007) WoLF PSORT: Protein localization predictor. Nucleic Acids Research. 35:W585-W587.

Hou B, Heidrich ES, Mehner-Breitfeld D, Brüser T (2018). The TatA component of the twin-arginine translocation system locally weakens the cytoplasmic membrane of Escherichia coli upon protein substrate binding. Journal of Biological Chemistry. 293:7592-7605.

Hu C, Lin SY, Chi WT, & Charng YY (2012) Recent gene duplication and subfunctionalization produced a mitochondrial GrpE, the nucleotide exchange factor of the Hsp70 complex, specialized in thermotolerance to chronic heat stress in Arabidopsis. Plant Physiology. 158:747-758.

Huang J, Hack E, Thornburg RW, & Myers AM (1990) A yeast mitochondrial leader peptide functions in vivo as a dual targeting signal for both chloroplasts and mitochondria. The Plant Cell. 2:1249-1260.

Jänsch L, Kruft V, Schmitz UK, & Braun HP (1998) Unique composition of the preprotein translocase of the outer mitochondrial membrane from plants. Journal of Biological Chemistry. 273: 17251-17257.

Jongbloed JD, Grieger U, Antelmann H, Hecker M, Nijland R, Bron S, & Van Dijl JM (2004). Two minimal Tat translocases in Bacillus. Molecular Microbiology. 54:1319-1325.

Jongbloed JD, van der Ploeg R & van Dijl JM (2006) Bifunctional TatA subunits in minimal Tat protein translocases. Trends in Microbiology. 14:2-4.

Joshi HJ, Hirsch-Hoffmann M, Baerenfaller K, Gruissem W, Baginsky S, Schmidt R, Schulze W X, Sun Q, van Wijk KJ, Egelhofer V, Wienkoop S, Weckwerth W, Bruley C, Rolland N., Toyoda T, Nakagami H, Jones AM, Briggs SP, Castleden I, Tanz SK, Millar AH, Heazlewood JL (2011) MASCP Gator: an aggregation portal for the visualization of Arabidopsis proteomics data. Plant Physiology. 155:259–270.

Kaddoum L, Magdeleine E, Waldo GS, Joly E and Cabantous S (2010) One-step split GFP staining for sensitive protein detection and localization in mammalian cells. Biotechniques. 49:727-8.

Kamiyama D, Sekine S, Barsi-Rhyne B, Hu J, Chen B, Gilbert LA, Ishikawa H, Leonetti MD, Marshall WF, Weissman JS and Huang B (2016) Versatile protein tagging in cells with split fluorescent protein. Nature Communications. 7:11046.

Keech O, Dizengremel P and Gardeström P (2005) Preparation of leaf mitochondria from Arabidopsis thaliana. Physiologia Plantarum. 124:403-409.

Kikuchi S, Asakura Y, Imai M, Nakahira Y, Kotani Y, Hashiguchi Y, Nakai Y, Takafuji K, Bédard J, Hirabayashi-Ishioka Y and Mori H (2018) A Ycf2-FtsHi heteromeric AAA-ATPase complex is required for chloroplast protein import. The Plant Cell. 30:2677-2703.

Klein TM, Arentzen R, Lewis PA, Fitzpatrick-McElligott S (1992) Transformation of microbes, plants and animals by particle bombardment. Nature Biotechnology. 10:286–291

Kmiec B, Teixeira PF, Glaser E (2014) Phenotypical consequences of expressing the dually targeted Presequence Protease, AtPreP, exclusively in mitochondria. Biochimie. 100:167-170

Ko K, & Cashmore AR (1989) Targeting of proteins to the thylakoid lumen by the bipartite transit peptide of the 33 kd oxygen-evolving protein. The EMBO Journal. 8:3187-3194.

Köhler RH, Zipfel WR, Webb WW and Hanson MR (1997) The green fluorescent protein as a marker to visualize plant mitochondria in vivo. The Plant Journal. 11:613-621.

Koncz C and Schell J (1986) The promoter of T L-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of Agrobacterium binary vector. Molecular and General Genetics. 204:383-396.

Krab IM, Parmeggiani A (1998) EF-Tu, a GTPase odyssey. Biochimica et Biophysica Acta. 1443:1-22.

Krause K, Krupinska K (2009) Nuclear regulators with a second home in organelles. Trends in Plant Science. 14:194–199.

Kuhlman P and Palmer JD (1995) Isolation, expression, and evolution of the gene encoding mitochondrial elongation factor Tu in Arabidopsis thaliana. Plant Molecular Biology. 29:1057-1070.

Kushnir S, Babiychuk E, Storozhenko S, Davey MW, Papenbrock J, De Rycke R, Engler G, Stephan UW, Lange H, Kispal G and Lill R (2001) A mutation of the mitochondrial ABC transporter Sta1 leads to dwarfism and chlorosis in the Arabidopsis mutant starik. The Plant Cell. 13:89-100.

Kwon KC, Verma D, Jin S, Singh, ND and Daniell H (2013) Release of proteins from intact chloroplasts induced by reactive oxygen species during biotic and abiotic stress. PLoS One. 8:e67106.

Lang BF, Gray MW, Burger G (1999) Mitochondrial Genome Evolution and the Origin of Eukaryotes. Annual Reviews of Genetics. 33:351–397.

Langner U, Baudisch B, Klösgen RB (2014) Organelle import of proteins with dual targeting properties into mitochondria and chloroplasts takes place by the general import pathways. Plant Signaling and Behavior. 9:e29301.

Lister R, Chew O, Rudhe C, Lee MN and Whelan J (2001) Arabidopsis thaliana ferrochelatase-I and-II are not imported into Arabidopsis mitochondria. FEBS Letters. 506:291-295.

Maarse AC, Van Loon AP, Riezman H, Gregor I, Schatz G and Grivell LA (1984) Subunit IV of yeast cytochrome c oxidase: cloning and nucleotide sequencing of the gene and partial amino acid sequencing of the mature protein. The EMBO Journal. 3:2831-2837.

Machettira AB, Gross LE, Sommer MS, Weis BL, Englich G, Tripp J and Schleiff E (2011) The localization of Tic20 proteins in Arabidopsis thaliana is not restricted to the inner envelope membrane of chloroplasts. Plant Molecular Biology. 77:381.

Mann GW, Calley PC, Joshi HJ and Heazlewood JL (2013) MASCP gator: an overview of the Arabidopsis proteomic aggregation portal. Frontiers in Plant Science. 4:411.

Margulis L (1981) Symbiosis in Cell Evolution: Life and its Environment on the Early Earth. Freeman Press, San Francisco.

Marques JP, Schattat MH, Hause G, Dudeck I and Klösgen RB (2004) In vivo transport of folded EGFP by the ΔpH/TAT-dependent pathway in chloroplasts of Arabidopsis thaliana. Journal of Experimental Botany. 55:1697-1706.

Martin T, Sharma R, Sippel C, Waegemann K, Soll J, & Vothknecht UC (2006) A protein kinase family in Arabidopsis phosphorylates chloroplast precursor proteins. Journal of Biological Chemistry. 281:40216-40223.

Martin W (2010) Evolutionary origins of metabolic compartmentalization in eukaryotes. Philosophical Transactions of Royal Society-B. 365:847–855.

Masuda T, Suzuki T, Shimada H, Ohta H and Takamiya K (2003) Subcellular localization of two types of ferrochelatase in cucumber. Planta. 217:602-609.

Mazzoleni M, Figuet S, Martin-Laffon J, Mininno M, Gilgen A, Leroux M, Brugière S, Tardif M, Alban C, Ravanel S (2015) Dual targeting of the protein methyltransferase PrmA contributes to both chloroplastic and mitochondrial ribosomal protein L11 methylation in Arabidopsis. Plant and Cell Physiology. 56:1697-1710.

McFadden GI (1999) Endosymbiosis and evolution of the plant cell. Curr Opin Plant Biol. 2:513–519.

Mitschke J, Fuss J, Blum T, Höglund A, Reski R, Kohlbacher O and Rensing SA (2009) Prediction of dual protein targeting to plant organelles. New Phytologist. 183:224-236.

Moseler A, Aller I, Wagner S, Nietzel T, Przybyla-Toscano J, Mühlenhoff U, Lill R, Berndt C, Rouhier N, Schwarzländer M and Meyer AJ (2015) The mitochondrial monothiol glutaredoxin S15 is essential for iron-sulfur protein maturation in Arabidopsis thaliana. Proceedings of the National Academy of Sciences. 112:13735-13740.

Motohashi R, Nagata N, Ito T, Takahashi S, Hobo T, Yoshida S, & Shinozaki (2001). An essential role of a TatC homologue of a ΔpH-dependent protein transporter in thylakoid membrane formation during chloroplast development in Arabidopsis thaliana. Proceedings of the National Academy of Sciences. 98:10499-10504.

Mueller SJ, Lang D, Hoernstein SN, Lang EG, Schuessele C, Schmidt A, Fluck M, Leisibach D, Niegl C, Zimmer AD and Schlosser A (2014) Quantitative analysis of the mitochondrial and plastid proteomes of the moss Physcomitrella patens reveals protein macrocompartmentation and microcompartmentation. Plant Physiology. 164:2081-2095.

Müller M, & Bernd Klösgen R (2005) The Tat pathway in bacteria and chloroplasts. Molecular Membrane Biology. 22:113-121.

Mulo P (2011) Chloroplast-targeted ferredoxin-NADP+ oxidoreductase (FNR): structure, function and location.Biochimica et Biophysica Acta (BBA)-Bioenergetics. 1807:927-934.

Murcha MW, Lister R, Ho AY and Whelan J (2003) Identification, expression, and import of components 17 and 23 of the inner mitochondrial membrane translocase from Arabidopsis. Plant Physiology. 131:1737–1747

Murcha MW, Elhafez D, Millar AH, & Whelan J (2004) The N-terminal extension of plant mitochondrial carrier proteins is removed by two-step processing: the first cleavage is by the mitochondrial processing peptidase. Journal of Molecular Biology. 344:443-454.

Murcha MW, Kmiec B, Kubiszewski-Jakubiak S, Teixeira PF, Glaser E, & Whelan J (2014). Protein import into plant mitochondria: signals, machinery, processing, and regulation. Journal of Experimental Botany. 65:6301-6335.

Nakai M (2018) New perspectives on chloroplast protein import. Plant and Cell Physiology. 59:1111-1119.

Nelson BK, Cai X and Nebenführ A (2007) A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. The Plant Journal. 51:1126-1136.

New CP, Ma Q, Dabney-Smith C (2018) Routing of thylakoid lumen proteins by the chloroplast twin arginine transport pathway. Photosynthesis Research. 138:289–301.

Nuñez PA, Soria M, & Farber MD (2012) The twin-arginine translocation pathway in α-proteobacteria is functionally preserved irrespective of genomic and regulatory divergence. PLoS One. 7:e33605.

Obara K, Sumi K, Fukuda H (2002) The Use of Multiple Transcription Starts Causes the Dual Targeting of Arabidopsis Putative Monodehydroascorbate Reductase to Both Mitochondria and Chloroplasts. Plant and Cell Physiology. 43:697–705

Oliver DJ, Neuburger M, Bourguignon J, & Douce R (1990) Interaction between the component enzymes of the glycine decarboxylase multienzyme complex. Plant Physiology. 94:833-839.

Ordon J, Gantner J, Kemna J, Schwalgun L, Reschke M, Streubel J, Jens Boch & Stuttmann J (2017) Generation of chromosomal deletions in dicotyledonous plants employing a user-friendly genome editing toolkit. The Plant Journal. 89:155-168.

Ottenhof HH, Ashurst JL, Whitney HM, Saldanha, SA, Schmitzberger F, Gweon HS, Blundell TL, Abell C and Smith AG (2004) Organisation of the pantothenate (vitamin B5) biosynthesis pathway in higher plants. The Plant Journal. 37:61-72.

Palmer Jeffrey D (1997) Organelle Genomes--Going, Going, Gone. Science. 275:790-790.

Papadakis AK and Roubelakis-Angelakis KA (2002) Oxidative stress could be responsible for the recalcitrance of plant protoplasts. Plant Physiology and Biochemistry. 40:549-559.

Park E, Lee HY, Woo J, Choi D and Dinesh-Kumar SP (2017) Spatiotemporal monitoring of Pseudomonas effectors via type III secretion using split fluorescent protein fragments. The Plant Cell. 29:1571-1584.

Petrů M, Wideman J, Moore K, Alcock F, Palmer T, Doležal P (2018) Evolution of mitochondrial TAT translocases illustrates the loss of bacterial protein transport machines in mitochondria. BMC Biology. 16:141.

Pruss GJ, Nester EW and Vance V (2008) Infiltration with Agrobacterium tumefaciens induces host defense and development-dependent responses in the infiltrated zone. Molecular Plant-Microbe Interactions. 21:1528-1538.

Rao RSP, Salvato F, Thal B, Eubel H, Thelen JJ and Møller IM (2017) The proteome of higher plant mitochondria. Mitochondrion. 33:22-37.

Reichel C, Mathur J, Eckes P, Langenkemper K, Koncz C, Schell J, Reiss B, Maas C (1996) Enhanced green fluorescence by the expression of an Aequorea victoria green fluorescent protein mutant in mono- and dicotyledonous plant cells. Proceeding of National Academy Sciences. 93:5888–5893.

Richardson LG, Singhal R, & Schnell DJ (2017) The integration of chloroplast protein targeting with plant developmental and stress responses. BMC Biology. 15:118.

Robinson C, & Ellis, RJ (1984) Transport of proteins into chloroplasts: partial purification of a chloroplast protease involved in the processing of imported precursor polypeptides. European Journal of Biochemistry. 142: 337-342.

Rödiger A, Baudisch B and Klösgen RB (2010) Simultaneous isolation of intact mitochondria and chloroplasts from a single pulping of plant tissue. Journal of Plant Physiology. 167:620-624.

Rödiger A, Baudisch B, Langner U, Klösgen RB (2011) Dual targeting of a mitochondrial protein: the case study of cytochrome c1. Molecular Plant. 4:679–687.

Rudhe C, Chew O, Whelan J and Glaser E (2002) A novel in vitro system for simultaneous import of precursor proteins into mitochondria and chloroplasts. The Plant Journal. 30:213-220.

Saisho D, Nakazono M, Lee KH, Tsutsumi N, Akita S and Hirai A (2001) The gene for alternative oxidase-2 (AOX2) from Arabidopsis thaliana consists of five exons unlike other AOX genes and is transcribed at an early stage during germination. Genes and Genetic Systems. 76:89-97.

Sapp, J., 2005. The prokaryote-eukaryote dichotomy: meanings and mythology. Microbiology and Molecular Biology Reviews. 69:292-305.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B and Tinevez JY (2012) Fiji: an open-source platform for biological-image analysis. Nature Methods. 9:676.

Schleiff E, & Becker T (2011).Common ground for protein translocation: access control for mitochondria and chloroplasts. Nature Reviews Molecular Cell Biology. 12:48.

Schleiff E, & Turnbull JL (1998) Functional and structural properties of the mitochondrial outer membrane receptor Tom20. Biochemistry. 37:13043-13051.

Senkler J, Senkler M, Eubel H, Hildebrandt T, Lengwenus C, Schertl P, Schwarzländer M, Wagner S, Wittig I and Braun HP (2017) The mitochondrial complexome of Arabidopsis thaliana. The Plant Journal. 89:1079-1092.

Serero A, Giglione C and Meinnel T (2001) Distinctive features of the two classes of eukaryotic peptide deformylases1. Journal of Molecular Biology. 314:695-708.

Sharma M, Bennewitz B and Klösgen RB (2018) Dual or Not Dual?—Comparative Analysis of Fluorescence Microscopy-Based Approaches to Study Organelle Targeting Specificity of Nuclear-Encoded Plant Proteins. Frontiers in Plant Science. 9:1350

Sharma M, Bennewitz B and Klösgen RB (2018) Rather rule than exception? How to evaluate the relevance of dual protein targeting to mitochondria and chloroplasts. Photosynthesis research. 138:335-343.

Sharma M, Kretschmer C, Lampe C, Stuttmann J, & Klosgen RB (2019) Determining targeting specificity of nuclear-encoded organelle proteins with the self-assembling split fluorescent protein toolkit. . Journal of Cell Science. 132:jcs230839.

Sheen J, Hwang S, Niwa Y, Kobayashi H, Galbraith DW (1995) Green fluorescent protein as a new vital marker in plant cells. Plant Journal. 8:777–784

Silva-Filho MC (1999) Translocation of a reporter protein into mitochondria is mediated by a chloroplast transit peptide and follows a normal import route. Journal of Plant Physiology. 154:51-54.

Small I, Peeters N, Legeai F, Lurin C (2004) Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics. 4:1581–1590

Staiger C, Hinneburg A and Klösgen RB (2009) Diversity in degrees of freedom of mitochondrial transit peptides. Molecular Biology and Evolution. 26:1773-1780.

Stanier RY and Niel CV (1962) The concept of a bacterium. Archives of Microbiology. 42:17-35.

Stengel A, Benz JP, Buchanan BB, Soll J, & Bölter B (2009) Preprotein import into chloroplasts via the Toc and Tic complexes is regulated by redox signals in Pisum sativum. Molecular Plant. 2:1181-1197.

Stuart RA (2002) Insertion of proteins into the inner membrane of mitochondria: the role of the Oxa1 complex. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research. 1592:79-87.

Sun Q, Zybailov B, Majeran W, Friso G, Olinares PDB, van Wijk KJ (2009) PPDB, the Plant Proteomics Database at Cornell. Nucleic Acids Research. 37:D969-74.

Sünkel S, Brennicke A, Knoop V (1994) RNA editing of a conserved reading frame in plant mitochondria increases its similarity to two overlapping reading frames in Escherichia coli. Molecular and General Genetics. 242:65-72.

Tabatabaei I, Dal Bosco C, Bednarska M, Ruf S, Meurer J and Bock R (2018) A highly efficient sulfadiazine selection system for the generation of transgenic plants and algae. Plant Biotechnology Journal. 17:638-649.

Tanz SK, Castleden I, Small ID and Millar AH (2013) Fluorescent protein tagging as a tool to define the subcellular distribution of proteins in plants. Frontiers in Plant Science. 4:214.

Tarasenko VI, Katyshev AI, Yakovleva TV, Garnik EY, Chernikova VV, Konstantinov YM, Koulintchenko MV (2016) RPOTmp, an Arabidopsis RNA polymerase with dual targeting, plays an important role in mitochondria, but not in chloroplasts. Journal of Experimental Botany. 67:5657-5669.

Taylor FJR (1976) Autogenous theories for the origin of eukaryotes. Taxon. 377-390.

Taylor L, Rudhe C, Hulett JM, Lithgow T, Glaser E, Day DA, Millar AH and Whelan J (2003) Environmental stresses inhibit and stimulate different protein import pathways in plant mitochondria. FEBS Letters. 547:125-130.

Teardo E, Carraretto L, De Bortoli S, Costa A, Behera S, Wagner R, Lo Schiavo F, Formentin E, Szabo I (2015) Alternative Splicing-Mediated Targeting of the Arabidopsis GLUTAMATE RECEPTOR3.5 to Mitochondria Affects Organelle Morphology. Plant Physiology. 167:216–227

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature. 408:796–815

Toyooka K, Moriyasu Y, Goto Y, Takeuchi M, Fukuda H, Matsuoka K (2006) Protein aggregates are transported to vacuoles by a macroautophagic mechanism in nutrient-starved plant cells. Autophagy. 2:96–106

Tsien RY (1998) The Green Florescent Protein. Annual Reviews of Biochemistry 67:509–544

Überlacker B and Werr W (1996) Vectors with rare-cutter restriction enzyme sites for expression of open reading frames in transgenic plants. Molecular Breeding. 2:293-295.

Ueda M, Nishikawa T, Fujimoto M, Takanashi H, Arimura SI, Tsutsumi N, Kadowaki KI (2008) Substitution of the gene for chloroplast RPS16 was assisted by generation of a dual targeting signal. Molecular Biology and Evolution. 25:1566-1575

Van Aken O, Zhang B, Carrie C, Uggalla V, Paynter E, Giraud E and Whelan J (2009) Defining the mitochondrial stress response in Arabidopsis thaliana. Molecular Plant. 2:1310-1324.

Van der Merwe JA and Dubery IA (2007) Expression of mitochondrial tatC in Nicotiana tabacum is responsive to benzothiadiazole and salicylic acid. Journal of Plant Physiology. 164:1231-1234.

Van Engelenburg SB and Palmer AE (2010) Imaging type-III secretion reveals dynamics and spatial segregation of Salmonella effectors. Nature Methods. 7:325.

Van Wijk KJ and Baginsky S (2011) Plastid proteomics in higher plants: current state and future goals. Plant Physiology. 155:1578-1588.

Voisine C, Craig EA, Zufall N, von Ahsen O, Pfanner N, & Voos W (1999) The protein import motor of mitochondria: unfolding and trapping of preproteins are distinct and separable functions of matrix Hsp70. Cell. 97:565-574.

Vojta L, Soll J, & Bölter B (2007) Protein transport in chloroplasts− targeting to the intermembrane space. The FEBS Journal. 274: 5043-5054.

Von Heijne G, Steppuhn J, Herrmann RG (1989) Domain structure of mitochondrial and chloroplast targeting peptides. European Journal of Biochemistry. 180:535-545.

Wagener N, Ackermann M, Funes S, Neupert W (2011) A pathway of protein translocation in mitochondria mediated by the AAA-ATPase Bcs1. Molecular cell. 44:191-202.

Wamboldt Y, Mohammed S, Elowsky C, Wittgren C, De Paula WB, Mackenzie SA (2009) Participation of leaky ribosome scanning in protein dual targeting by alternative translation initiation in higher plants. Plant Cell. 21:157–167.

Wang W, Vinocur B, Shoseyov O and Altman A (2004) Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends in Plant Science. 9:244-252.

Wang Z, Wang F, Hong Y, Huang J, Shi H & Zhu JK (2016) Two chloroplast proteins suppress drought resistance by affecting ROS production in guard cells. Plant Physiology. 172:2491-2503.

Watanabe N, Che FS, Iwano M, Takayama S, Yoshida S, Isogai A (2001) Dual Targeting of Spinach Protoporphyrinogen Oxidase II to Mitochondria and Chloroplasts by Alternative Use of Two In-frame Initiation Codons. Journal of Biological Chemistry. 276:20474–20481.

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.

Wetzel CM, Jiang CZ, Meehan LJ, Voytas DF and Rodermel SR (1994) Nuclear-organelle interactions: the immutans variegation mutant of Arabidopsis is plastid autonomous and impaired in carotenoid biosynthesis. The Plant Journal. 6:161-175.

Wroblewski T, Tomczak A and Michelmore R (2005) Optimization of Agrobacterium-mediated transient assays of gene expression in lettuce, tomato and Arabidopsis. Plant Biotechnology Journal. 3:259-273.

Wu FH, Shen SC, Lee LY, Lee SH, Chan MT and Lin CS (2009) Tape-Arabidopsis Sandwich-a simpler Arabidopsis protoplast isolation method. Plant Methods. 5:16.

Xiang C, Han P, Lutziger I, Wang K and Oliver DJ (1999) A mini binary vector series for plant transformation. Plant Molecular Biology. 40:711-717.

Xiong E, Zheng C, Wu X, Wang W (2016) Protein Subcellular Location: The Gap Between Prediction and Experimentation. Plant Molecular Biology Reporter. 34:52–61.

Xu YZ, De la Rosa Santamaria R, Virdi KS, Arrieta-Montiel MP, Razvi F, Li S, Ren G, Yu B, Alexander D, Guo L and Feng X (2012) The chloroplast triggers developmental reprogramming when MUTS HOMOLOG1 is suppressed in plants. Plant Physiology. 159:710-720.

Xu L, Carrie C, Law SR, Murcha MW and Whelan J (2013) Acquisition, conservation, and loss of dual-targeted proteins in land plants. Plant Physiology. 161:644-662.

Yao N, Greenberg JT (2006) Arabidopsis ACCELERATED CELL DEATH2 modulates programmed cell death. The Plant Cell. 18:397-411.

Zhang B, Van Aken O, Thatcher L, De Clercq I, Duncan O, Law SR, Cazzonelli C (2014) The mitochondrial outer membrane AAA ATP ase At OM 66 affects cell death and pathogen resistance in A rabidopsis thaliana. The Plant Journal. 80:709-727.

Zhang H, Whitelegge JP and Cramer WA, (2001) Ferredoxin: NADP+ oxidoreductase is a subunit of the chloroplast cytochrome b6f Complex. Journal of Biological Chemistry. 276:38159-38165.

Zhang XP and Glaser E (2002) Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone. Trends in Plant Science. 7:14-21.

Notes

This page is intentionally left blan

Appendix

Appendix-I Gene sequences

Name	Sequence (5′ – 3′)
ATPS 1-100 AA (*Arabidopsis thaliana*)	ATGGCGAGTCGGCGAATCTTATCATCGCTTCTCCGTTCTTCTTCGAGTAGATCTAC TTCTAAATCCTCCTTGATCGGGAGCCGAAACCCGAGGCTTTTATCCCCCGGTCCCG CGCATGGAGCCGCTCCATGTGGGACTCTCCTTGGCCGAGTCGCCGAGTATTCGACT TCTTCTCCGGCTAATTCGGCTGCACCATCTTCTGCTCCTGCTAAAGATGAGGGAAA GAAGACCTATGATTACGGTGGCAAAGGCGCGATCGGGCGTGTGTGTCAGGTTATTG GTGCCATTGTTGATGTGAGA
CoxIV 1-29 AA (*Saccharomyces cerevisiae*)	ATGCTTTCACTACGTCAATCTATAAGATTTTTCAAGCCAGCCACAAGAACTTTGTG TAGCTCTAGATATCTGCTTCAGCAAAAACCC
Ef-Tu 1-100 AA (*Arabidopsis thaliana*)	ATGGCGTCCGTTGTTCTTCGAAACCCTAGCTCGAAGCGCCTTGTTCCATTCTCTTC CCAGATCTACTCTCGCTGTGGTGCTTCCGTTACTTCCTCTTACTCGATCTCTCATT CTATCGGTGGAGATGATCTCTCTTCCTCTACCTTCGGAACCTCCTCCTTCTGGAGA TCCATGGCCACTTTTACTCGAAATAAACCTCATGTAAATGTTGGAACTATTGGGCA TGTTGATCATGGCAAGACCACTTTAACTGCTGCAATCACAAAGGTTCTTGCTGAGG AGGGCAAAGCTAAAGCTATT
eYFP	ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCT GGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATG CCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTG CCCTGGCCCACCCTCGTGACCACCTTCGGCTACGGCCTGCAGTGCTTCGCCCGCTA CCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACG TCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAG GTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTT CAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACA ACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATC CGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACAC CCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGT CCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAA
FNR 1-55 (*Spinacia oleracea*)	ATGACCACCGCTGTCACCGCCGCTGTTTCTTTCCCCTCTACCAAAACCACCTCTCT CTCCGCCCGAAGCTCCTCCGTCATTTCCCCTGACAAAATCAGCTACAAAAAGGTTC CTTTGTACTACAGGAATGTATCTGCAACTGGGAAAATGGGACCCATCAGGGCC
GCS 1-100 AA (*Arabidopsis thaliana*)	ATGGCACTAAGAATGTGGGCTTCTTCTACAGCAAACGCTCTCAAGCTTTCTTCTTC TGTTTCCAAGTCTCATCTCTCTCCTTTCTCCTTCTCTAGATGCTTCTCCACAGTTT TGGAGGGTTTGAAGTATGCAAATTCACATGAGTGGGTTAAACATGAAGGCTCTGTT GCCACCATTGGCATCACTGCCCATGCTCAGGACCATTTAGGTGAAGTGGTGTTTGT TGAACTGCCAGAGGACAATACTTCAGTGAGCAAAGAGAAAAGCTTTGGAGCAGTGG AGAGTGTGAAGGCAACAAGT
GFP1-10 (Kamiyama et al., 2015)	ATGTCCAAAGGAGAAGAACTGTTTACCGGTGTTGTGCCAATTTTGGTTGAACTCGA TGGTGATGTCAACGGACATAAGTTCTCAGTGAGAGGCGAAGGAGAAGGTGACGCCA CCATTGGAAAATTGACTCTTAAATTCATCTGTACTACTGGTAAACTTCCTGTACCA TGGCCGACTCTCGTAACAACGCTTACGTACGGAGTTCAGTGCTTTTCGAGATACCC AGACCATATGAAAAGACATGACTTTTTTAAGTCGGCTATGCCTGAAGGTTACGTGC AAGAAAGAACAATTTCGTTCAAAGATGATGGAAAATATAAAACTAGAGCAGTTGTT AAATTTGAAGGAGATACTTTGGTTAACCGCATTGAACTGAAAGGAACAGATTTTAA AGAAGATGGTAATATTCTTGGACACAAACTCGAATACAATTTTAATAGTCATAACG TATACATCACTGCTGATAAGCAAAAGAACGGAATTAAAGCGAATTTCACAGTACGC CATAATGTAGAAGATGGCAGTGTTCAACTTGCCGACCATTACCAACAAAACACCCC TATTGGAGACGGTCCGGTACTTCTTCCTGATAATCACTACCTCTCAACACAAACAG TCCTGAGCAAAGATCCAAATGAAAAAGGAACAGGTGGCGGCGGAAGTTAG
GFP11X1	ATGGGCGGCAAATTCATGCGTGACCACATGGTCCTTCATGAGTATGTAAATGCTGC TGGGATTACATAA
GFP11X3	ATGGGCGGCAAATTCATGCGTGACCACATGGTCCTTCATGAGTATGTAAATGCTGC TGGGATTACAGGTGGCTCTGGAGGTAGAGATCATATGGTTCTCCACGAATACGTTA

	ACGCCGCAGGCATCACTGGCGGTAGTGGAGGACGCGACCATATGGTACTACATGAA TATGTCAATGCAGCCGGAATAACCGGATAA
GFP11x7 (Kamiyama et al., 2015)	ATGGGCGGCAAATTCATGCGTGACCACATGGTCCTTCATGAGTATGTAAATGCTGC TGGGATTACAGGTGGCTCTGGAGGTAGAGATCATATGGTTCTCCACGAATACGTTA ACGCCGCAGGCATCACTGGCGGTAGTGGAGGACGCGACCATATGGTACTACATGAA TATGTCAATGCAGCCGGAATAACCGGAGGGTCCGGAGGCCGGGATCACATGGTGCT GCATGAGTATGTGAACGCGGCGGGTATAACTGGTGGGTCGGGCGGACGAGATCATA TGGTGCTTCACGAATACGTAAACGCAGCTGGCATTACTGGCGGATCAGGTGGCAGG GATCACATGGTACTCCATGAGTACGTGAACGCTGCTGGAATCACAGGCGGTAGCGG CGGTCGGGACCATATGGTCCTGCACGAATATGTCAATGCTGCCGGTATCACCTAA
GrpE 1-100 AA (*Arabidopsis thaliana*)	ATGTTGGTGTCTAGGGTTTTATCAAGAGTTTCTCGCAGCGCAGGCTTACGCTCGTC TTTCTCCTCCGTCGTTACTCCGAAGAGGAATCAGATTCCGATTGTTGCAAGCCGAT TTCACTCCCTCGTTCACGGAACCCCCAACAAGCTTGTTGCAGTTCCAGTGTCGCTC CGAAACCATGGAACTCTAGATTTGAATGTTCTTCAAAGGTTCGGCTTTTTTTCCTC TTCCTCAGCTGAACCGAAGGGAAATGAGAGTAACACCGAAGTACCAAAGACCGGCG AAACTTCTGAGAATGTGGAG
Gtred 1-100 AA (*Arabidopsis thaliana*)	TGGCGGCTTCTTTATCGAGCAGACTTATAAAAGGAATCGCTAATCTCAAAGCTGTT CGTTCTAGCAGATTGACGTCTGCATCAGTCTACCAAAATGGGATGATGAGATTTTC CTCAACAGTGCCAAGTGATTCAGATACACATGATGATTTCAAGCCTACACAAAAAG TCCCTCCCGATTCTACGGACTCACTTAAAGATATCGTTGAGAATGATGTGAAGGAT AATCCTGTTATGATCTACATGAAAGGTGTCCCTGAATCTCCTCAGTGTGGGTTTAG CTCACTAGCCGTCAGAGTTT
mtRi 1-100 AA (*Solanum tuberosum*)	ATGCTTCGAGTAGCAGGTAGAAGGCTTTCTTCTTCAGCCGCTAGATCTTCATCTAC CTTCTTTACAAGAAGCTCTTTCACCGTTACCGATGATTCGTCTCCGGCAAGATCTC CTTCTCCGTCACTCACCTCTTCGTTTCTCGATCAAATCAGAGGTTTCTCATCTAAT TCGGTTTCTCCCGCACATCAGTTGGGTTTAGTCTCAGATCTTCCAGCCACAGTGGC TGCTATTAAGAATCCCAGTTCAAAAATTGTATATGATGACTCCAACCATGAGCGTT ATCCACCTGGAGATCCTAGC
PDF 1-100 AA (*Arabidopsis thaliana*)	ATGGCCGTCTGTAACTGCTTCCTCCAAGCTCCACCACTCTCTCGCATACTCCTACC GGTTTTATCTCGCCGAGCTACTACTCTCTCTGCCGGTTATGGTCGGCTCAAATCCA CGGTCACGTTCTGCTCTACTGTGAACCGGACCAGTCCTCTGACATCTTCGGTCCGC GCAGAAGTAAAGCGCGTCTCGCGTAAAGACGATAAAGTAGCTTCAGCTACTGATGT TCAATTTGAGACGCCTTTGAAGATTGTTGAGTATCCGGATCCTATACTACGGGCTA AGAACAAGAGGATTGATATT
PTOX (*Arabidopsis thaliana*)	ATGGCGGCGATTTCAGGCATCTCCTCTGGTACGTTGACGATTTCACGGCCTTTGGT TACTCTTCGACGCTCTAGAGCCGCCGTTTCGTACAGCTCCTCTCACCGATTGCTTC ATCATCTTCCTCTCTCTTCTCGTCGTCTGCTATTAAGGAACAATCATCGAGTCCAA GCAACGATTTTGCAAGACGATGAAGAGAAAGTGGTGGTGGAGGAATCGTTTAAAGC CGAGACTTCTACTGGTACAGAACCACTTGAGGAGCCAAATATGAGTTCTTCTTCAA CTAGTGCTTTTGAGACATGGATCATCAAGCTTGAGCAAGGAGTGAATGTTTTCCTT ACAGACTCGGTTATTAAGATACTTGACACTTTGTATCGTGACCGAACATATGCAAG GTTCTTTGTTCTTGAGACAATTGCTAGAGTGCCTTATTTTGCGTTTATGTCTGTGC TACATATGTATGAGACCTTTGGTTGGTGGAGGAGAGCAGATTATTTGAAAGTACAC TTTGCTGAGAGCTGGAATGAAATGCATCACTTGCTCATAATGGAAGAATTGGGTGG AAATTCTTGGTGGTTTGATCGTTTTCTGGCTCAGCACATAGCAACCTTCTACTACT TCATGACAGTGTTCTTGTATATCTTAAGCCCTAGAATGGCATATCACTTTTCGGAA TGTGTGGAGAGTCATGCATATGAGACTTATGATAAATTTCTCAAGGCCAGTGGAGA GGAGTTGAAGAATATGCCTGCACCGGATATCGCAGTAAAATACTATACGGGAGGTG ACTTGTACTTATTTGATGAGTTCCAAACATCAAGAACTCCCAATACTCGAAGACCA GTAATAGAAAATCTATACGATGTGTTTGTGAACATAAGAGATGATGAAGCAGAACA CTGCAAGACAATGAGAGCTTGTCAGACTCTAGGCAGTCTGCGTTCTCCACACTCCA TTTTAGAAGATGATGATACTGAAGAAGAATCAGGGTGTGTTGTTCCTGAGGAGGCT CATTGCGAAGGTATTGTAGACTGCCTCAAGAAATCCATTACAAGTTAA
RbcS 1-79 AA (*Nicotiana sylvestris*)	ATGGCTTCCTCAGTTCTTTCCTCTGCAGCAGTTGCCACCCGCAGCAATGTTGCTCA AGCTAACATGGTTGCACCTTTCACTGGCCTTAAGTCAGCTGCCTCATTCCCTGTTT CAAGGAAGCAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGCAA TGCATGCAGGTGTGGCCACCAATTAACAAGAAGAAGTACGAGACTCTCTCATACCT TCCTGATTTGAGC
TatA (*Arabidopsis thaliana*)	ATGGCGACATCTGTTGCGACTCTATCTTCTCCACCACCAGTATCTCTACCTCTCCT CTCATCATCTCGCTCTTCCTTCTTCTCCAATTGCTTCACAGTCACGACCCGACCAA

	ACACTCGTTCTCTGGTAGCGATTGGACGCAGAATCCGACAAGAACCAACGAGAAAG CCACTGACTTGTAATGCTTTGTTTGGTCTCGGTGTGCCTGAGCTTGCTGTGATTGC TGGCGTCGCTGCGTTGCTTTTTGGACCTAAGAAACTTCCTGAGATTGGCAAGAGTA TTGGCAAAACCGTCAAGAGCTTTCAACAGGCTGCAAAAGAGTTCGAGTCTGAGCTT AAGACAGAACCCGAAGAGTCTGTTGCAGAATCATCTCAGGTAGCAACGAGTAACAA AGAAGAGGAGAAAAAAACTGAGGTTTCTTCAAGCTCAAAGGAGAATGTA
TatA 1-100 AA (*Pisum sativum*)	ATGGAGATAACACTTTCCATTTCTTCATCTTCAGTGATTCCAACTAGACTACCAAA CTCTTCATGCTATTCCAATTTATCCTTCTTATCTTCCAACTCTAACACTTCCTCAC TACTCTTGAAGAAAGCCAGAATCAAAACCAGAACAACAAAGGGTTTCACCTGCAAT GCCTTCTTCGGTCTAGGCGTGCCTGAGCTTGTTGTTATTGCGGGAGTCGCTGCTCT TGTTTTCGGTCCCAAGAAATTGCCCGAAGTCGGTCGCAGCATCGGCCAAACTGTCA AAAGCTTTCAACAGGCAGCA
TatA gene after deletion of fragment via Crispr/Cas9 (Arabidopsis thaliana)	ATTCCCTTGCAAAACAACAACCGGAAAACCAAATTTTGATTACAAAAGAAATTAGG GTTCACATTCAAAACAAAACGCTGATTCGGTACTTACAAACAATAGGAAAAATTCG GCGATGGAGATTTGGAACGGCGGCGGTAACGGCGGATTGTAACGGCGGGTGGTGAA GAAGCAGAGAGGAGAGAGATCCGGCGATGGAGATTTGGAAAGATTCGGCGATGGCG ATGGAGATTTGCTCGGCGAGAAAGGATTTCGATCGAAATCGTCGTTCGTCAGTTTA GGATTTCGAAAATGACGGAAGTTCATATTAATAATTAGGGTAAATAATTCGGTTTT AGGGGAAAGCCCATTAGAACTAGGTTGGATCCAGTCCTATCCGCAGATTTGCGGTT GGATCTTGTAAATTTTTTAAAGGCATAAGATTTTTTAATAATTTATTATGAGTTTT TATATAAAATTAAATGAAAACTATATTAGATAAATGTGACAGAATATCAATTTTTC ATTTGAAACCCAGCTAAAAGCCTAAAACAAATATTATTTTATTTAATGGATTATAT TGGGTAAAACCAAGAAAAAGAAAACACAGATAACATTTAGAACCAGAAAAAAAAAA AAAAGTAGCAGCAATGGCGACATCTGTTGCGACTCTAACTTCATGTTTCAGGTGCT TAGTTTTTCTATGAGTTTATGTTTAGATTTGTAAGGAATTGATAAGAATACTAAGA AAAACTTGGAAACTGAGTGGTTTATATAACTTTGCTGTATGACCGTAGATGAAGAA TAAATCATTCATTTGCTTGTAAACATTAAGTTTCATTGATAACTTTGCTGTGTATT ATAAATCAAAGCATGTTATGTTAGCTCAAACGGGGAAATTTATTCAATTCATGTGA ATAGCTTGTTAGTTGTAGTTGGGATTGATGTAAAATGAATGTGTATACTTTTTGTG GATTTGTGAGGAACTGATAAGAACACTCAGAAAGATTAGTTACTGAGTGGGGACAA CTTTACCTTATAGAGTATATAGATCAAGAATAAAACATTTTATTTGCTTGTAACAG TAAGTTCATTCATGAAATGTCGATCAAGCTTGTTGTTATGTTATGTTATGTTAAAT CAATAATAAAATTTTGCTATATAGTATAGACCAAGCATGTTGTTATGTTAGGTTAA TCTGGCAATGCAATCTCTATTCAAAAACCGATCACAGAGGTAGCTCGTGATTTTAT TATTGGAGTGATATTAGCTCCGAGAAAACTTCTGGTTATTGTTATGGCTCTCATGG TTGTCTGTTTGTTGTATGTTTTTCTGTGTTCCTATTATCGTTTTACTTGTATTGAT GAAGCTTTGTATGGTTTGGTATAACGCAGGCTGCAAAAGAGTTCGAGTCTGAGCTT AAGACAGAACCCGAAGAGTCTGTTGCAGAATCATCTCAGGTAGCAACGAGTAACAA AGAAGAGGAGAAAAAAACTGAGGTTTCTTCAAGCTCAAAGGAGAATGTATGAAGAT GAGAGACTTCAGTGTTTGCTTTTCCTTGTAGCTACAACAATTTTATGTCATGAAGA GTTTAGATCAAATATTATTGTCTGGAGGAAACAAAATAGCCTTCCAATGAAAAGTA TAAATTAATTTGTGTTTGATGAATCTAGAAGGCTCTTTCGAGT
TatA gene with regulatory elements (*Arabidopsis thaliana*)	ATTCCCTTGCAAAACAACAACCGGAAAACCAAATTTTGATTACAAAAGAAATTAGG GTTCACATTCAAAACAAAACGCTGATTCGGTACTTACAAACAATAGGAAAAATTCG GCGATGGAGATTTGGAACGGCGGCGGTAACGGCGGATTGTAACGGCGGGTGGTGAA GAAGCAGAGAGGAGAGAGATCCGGCGATGGAGATTTGGAAAGATTCGGCGATGGCG ATGGAGATTTGCTCGGCGAGAAAGGATTTCGATCGAAATCGTCGTTCGTCAGTTTA GGATTTCGAAAATGACGGAAGTTCATATTAATAATTAGGGTAAATAATTCGGTTTT AGGGGAAAGCCCATTAGAACTAGGTTGGATCCAGTCCTATCCGCAGATTTGCGGTT GGATCTTGTAAATTTTTTAAAGGCATAAGATTTTTTAATAATTTATTATGAGTTTT TATATAAAATTAAATGAAAACTATATTAGATAAATGTGACAGAATATCAATTTTTC ATTTGAAACCCAGCTAAAAGCCTAAAACAAATATTATTTTATTTAATGGATTATAT TGGGTAAAACCAAGAAAAAGAAAACACAGATAACATTTAGAACCAGAAAAAAAAAA AAAAGTAGCAGCAATGGCGACATCTGTTGCGACTCTATCTTCTCCACCACCAGTAT CTCTACCTCTCCTCTCATCATCTCGCTCTTCCTTCTTCTCCAATTGCTTCACAGTC ACGACCCGACCAAACACTCGTTCTCTGGTAGCGATTGGACGCAGAATCCGACAAGA ACCAACGAGAAAGCCACTGACTTGTAATGCTTTGTTTGGTCTCGGTGTGCCTGAGC TTGCTGTGATTGCTGGCGTCGCTGCGTTGCTTTTTGGACCTAAGAAACTTCCTGAG ATTGGCAAGAGTATTGGCAAAACCGTCAAGAGCTTTCAACAGGTTTGCTACATAGA AATTTTGTTGGGATCTTTATGAAATTCTTTGAGTTCTGCTTTGTTATTGTATGAGT TGTGTCTGGATTTGTGAGGAATTGATAACAAGTTTCAAAAAATTGTGCAATTCTGG

	AGATTCTTGTGGTAAGAAGAAATGTGATTTTGGGACCTTTCTGTTTCAGATGCTTA GTTATGTTTGGATTTGTGTGGAATTGAGTGGTTATATAACTTTGCTATATAGAACA TCTGCTTGTAAACTTTGCTGTATAGGATCATAACAAATTTGATAGAATATGGATGT GATAATCTTCTTATGCCAAGTCATGTGAATAGCTTGTTGATTGTTGTTGGAATTGA TGTAAAAGGTCAAAATGAGGAAACTTCATGTTTCAGGTGCTTAGTTTTTCTATGAG TTTATGTTTAGATTTGTAAGGAATTGATAAGAATACTAAGAAAAACTTGGAAACTG AGTGGTTTATATAACTTTGCTGTATGACCGTAGATGAAGAATAAATCATTCATTTG CTTGTAAACATTAAGTTTCATTGATAACTTTGCTGTGTATTATAAATCAAAGCATG TTATGTTAGCTCAAACGGGGAAATTTATTCAATTCATGTGAATAGCTTGTTAGTTG TAGTTGGGATTGATGTAAAATGAATGTGTATACTTTTTGTGGATTTGTGAGGAACT GATAAGAACACTCAGAAAGATTAGTTACTGAGTGGGGACAACTTTACCTTATAGAG TATATAGATCAAGAATAAAACATTTTATTTGCTTGTAACAGTAAGTTCATTCATGA AATGTCGATCAAGCTTGTTGTTATGTTATGTTATGTTAAATCAATAATAAAATTTT GCTATATAGTATAGACCAAGCATGTTGTTATGTTAGGTTAATCTGGCAATGCAATC TCTATTCAAAAACCGATCACAGAGGTAGCTCGTGATTTTATTATTGGAGTGATATT AGCTCCGAGAAAACTTCTGGTTATTGTTATGGCTCTCATGGTTGTCTGTTTGTTGT ATGTTTTTCTGTGTTCCTATTATCGTTTTACTTGTATTGATGAAGCTTTGTATGGT TTGGTATAACGCAGGCTGCAAAAGAGTTCGAGTCTGAGCTTAAGACAGAACCCGAA GAGTCTGTTGCAGAATCATCTCAGGTAGCAACGAGTAACAAAGAAGAGGAGAAAAA AACTGAGGTTTCTTCAAGCTCAAAGGAGAATGTATGAAGATGAGAGACTTCAGTGT TTGCTTTTCCTTGTAGCTACAACAATTTTATGTCATGAAGAGTTTAGATCAAATAT TATTGTCTGGAGGAAACAAAATAGCCTTCCAATGAAAAGTATAAATTAATTTGTGT TTGATGAATCTAGAAGGCTCTTTCGAGTAGAAGAAGCAAAGAACTTGCACAAACAA ATTAACTCTAGTTAAGACCAAAGACCAAAGAAATACAGAGAAC
TyrRS 1-91 AA (*Arabidopsis thaliana*)	ATGGCATATGCAACAGGAATAACGTTTGCCTCAAGGAGTATTTTGCCTATTTGTTC CAGAACCTTCTTATCTCCCTTGCGTGTCGCTTCTTTACTCGTCTTTCCTGAGAAAT CATCTGCAACTTTCTTCAGAAGGGTCCAAGTTCCTCACCTTTTTTCCACTTCTACT ACTACTCTGTTCTCTTCTGTTAAGTGTTCAATTCATTCTACTTCATCTCCTGAGAC AGAGAATCAAGCTGTTTTTCGCCCTAATGTAGTCGATATACTTGAAGAA
YFP1-10	ATGTCCAAAGGAGAAGAACTGTTTACCGGTGTTGTGCCAATTTTGGTTGAACTCGA TGGTGATGTCAACGGACATAAGTTCTCAGTGAGAGGCGAAGGAGAAGGTGACGCCA CCATTGGAAAATTGACTCTTAAATTCATCTGTACTACTGGTAAACTTCCTGTACCA TGGCCGACTCTCGTAACAACGCTTACGTACGGAGTTCAGTGCTTTTCGAGATACCC AGACCATATGAAAAGACATGACTTTTTTAAGTCGGCTATGCCTGAAGGTTACGTGC AAGAAAGAACAATTTCGTTCAAAGATGATGGAAAATATAAAACTAGAGCAGTTGTT AAATTTGAAGGAGATACTTTGGTTAACCGCATTGAACTGAAAGGAACAGATTTTAA AGAAGATGGTAATATTCTTGGACACAAACTCGAATACAATTTTAATAGTCATAACG TATACATCACTGCTGATAAGCAAAAGAACGGAATTAAAGCGAATTTCACAGTACGC CATAATGTAGAAGATGGCAGTGTTCAACTTGCCGACCATTACCAACAAAACACCCC TATTGGAGACGGTCCGGTACTTCTTCCTGATAATCACTACCTCTCATATCAAACAG TCCTGAGCAAAGATCCAAATGAAAAAGGAACAGGTGGCGGCGGAAGTTAG

Appendix-II Step-by-step protocols

(A) Tape-Sandwich method for protoplast isolation

Materials Required

1. Enzyme solution- (pH 5.7; adjust with KOH)

Component	Final Conc.	Stock Conc.	Amount added (50ml)
Cellulase (R10)	1% w/v		0.5 g
Macerozyme (R10)	0.25% w/v		0.25 g
Mannitol	0.4 M	1.0 M	20 ml
$CaCl_2$	10 mM	1.0 M	0.5 ml
KCl	20 mM	1.0 M	1 ml
BSA	0.1%	1 %	5 ml
MES	20 mM	0.1 M	10 ml
ddWater			Upto 50 ml

2. W5 Solution- (pH 5.7; adjust with KOH)

Component	Final Conc.	Stock Conc.	Amount added (150ml)
NaCl	154 mM	2.5 M	9.24 ml
$CaCl_2$	125 mM	1.0 M	18.75 ml
KCl	5 mM	1.0 M	0.75 ml
Glucose	5 mM	1.0 M	0.75 ml
MES	2 mM	0.1 M	3 ml
ddWater			Upto 150 ml

3. Modified Mg Solution- (pH 5.7; adjust with KOH)

Component	Final Conc.	Stock Conc.	Amount added (50ml)
Mannitol	0.4 M	1.0 M	20 ml
$MgCl_2$	15 mM	0.1 M	7.5 ml
MES	4 mM	0.1 M	2 ml
ddWater		0.2	Upto 50 ml

4. *Arabidopsis* plants (3-6 week old)

Protocol-

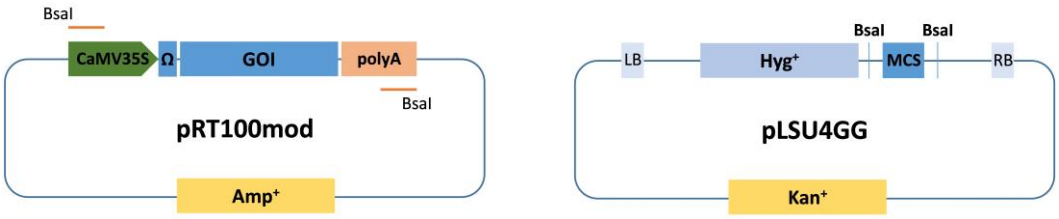
1. Cut leaves (3 to 5-week-old plants grown under optimal light (ca. 150 µE m^{-2} s^{-1}).
2. The upper epidermal surface was stabilized by affixing a strip of normal tape (coloured ones) while the lower epidermal surface was affixed to a strip of Scotch tape.
3. The Scotch tape was then carefully pulled away from the normal tape, peeling away the lower epidermal surface cell layer.
4. The peeled leaves (7 to10 optimal-light-growth leaves, about 1-2 g, up to 5 g), still adhering to the Time tape, were transferred to a Petri dish containing 5-10 mL of enzyme solution.

5. The leaves were gently shaken (40 rpm on a platform shaker) in light for 20 to 60 min until the protoplasts were released into the solution.
6. Transfer protoplast solution in a test-tube.
7. The protoplasts were centrifuged at 100 × *g* for 3 min in rotor, washed twice with
8. 12.5 mL of pre-chilled modified W5 solution and incubated on ice for 30 min.
9. During the incubation period, protoplasts can be counted using a haemocytometer under a light microscope
10. The protoplasts were then centrifuged and resuspended in modified MMg solution to a final concentration of 2 to 5 × 10^5 cells/mL (~500µl protoplast+ 500µl MMg).
11. Wash with 3 ml of W5 solution. Centrifuge at 100xg for 1 min. Remove 2 ml of supernatant.
12. Incubate protoplasts in MMg solution (1 ml) overnight in continuous light.

Mitotraker Staining

1. 0.01 µM Mitotracker-orange was added to protoplast in 2 ml Eppendorf tube.
2. Incubate for 15 min. in dark.
3. Centrifuge for 1 min at 100xg.
4. Wash with equal volume of W5 solution twice.
5. Resuspend in 1 ml W5 solution, observe under microscope.

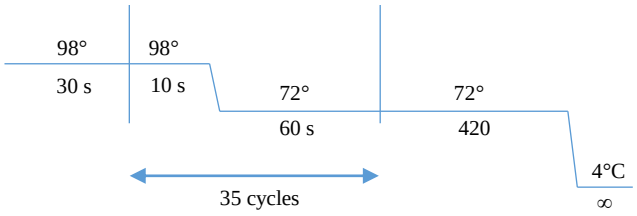
(B) Cloning of fragments from pRT100mod vectors to pLSU4GG



PCR Reaction (to amplify fragment of interest)

Components	Amount
GC Buffer (Phusion)	4.0 µl
dNTP	2.0 µl
Primer (GG pRT For)	0.4 µl
Primer (GG pRT Rev)	0.4 µl
Template (pLSU4GG)	X µl
Phusion	0.2 µl
ddH_2O	Up to 20 µl

Reaction Cycle (2-step PCR)

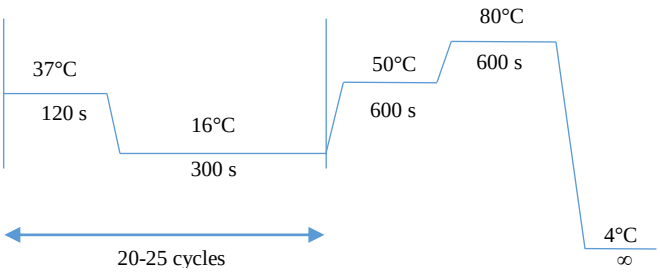


Note- Purify the PCR product from Agarose gel with gel extraction kit.

Golden Gate Digestion/Ligation Reaction

Components	Amount
Vector	~50 ng
PCR Product	~50-100 ng
Ligase Buffer	1.5 µl
10× BSA	1.5 µl
BsaI Enzyme	0.75 µl
T4 DNA Ligase	0.75 µl
ddH2O	Up to 15 µl

Reaction Cycle



Transform *E. coli* DH5α, electro-competent cells with 1.5 µl of reaction mix. (Selection- *Kanamycin*)

Acknowledgements-

Science is not a one-man show and so was this dissertation. I would like to take this opportunity to thank those people who directly or indirectly contributed to the successful completion of this dissertation. The first person in this acknowledgment list is my "Doktorvater" Prof. Dr. Ralf Bernd Klösgen (aka Rabe). Thanks to him for accepting me as a doctoral student, for his supervision, for giving me the freedom to design and implement my own ideas, for keeping his doors open for me and for all his support during my stay in the institution. It was a great experience working with him.

Institution Colleagues

I would like to acknowledge the constant support and care by Dr. Bationa Bennewitz. Thanks to Dr. Bennewitz for taking efforts to strengthen the dual targeting group, for proofreading all my writings and for her constructive criticism during the work. A big thanks to all doctoral colleagues namely Mr. Matthias Reimers, for all his small but 'Groß' help, for making my stay more comfortable in the lab and for last minute proofreading of this dissertation; Ms. Sarah Zinecker and Ms. Jana Kungel for their chilling attitudes and great companionship throughout my lab tenure; Ms. Julia Dittmar for being a nice colleague and for her occasional support in the lab; Mrs. Jessica Lee Erickson for her insights and small talks while passing by the office. I would also like to acknowledge Dr. Mario Jakob for his suggestions in lab seminars and the technicians Ms. Ulrike Gosdzenski and Mrs. Birgit Kretschmann for providing helping hands in my work. Also, thanks to Mr. Siegfried Platzer for taking care of the plants in the greenhouse.

Collaborators

Thanks to Dr. Martin Schattat for providing the construct (pLSU4GG), for his novel ideas about microscopy and for the deep discussions about science. In addition, I acknowledge Dr. Johannes Stuttmann for providing various DNA constructs and for introducing me to the magical world of Golden Gate cloning. I really appreciate their openness about science and collaborations. Thanks to Dr. Gerd Hause for performing electron microscopy of the mutants on a very short notice. I also acknowledge Prof. Elzbieta Glaser (Sweden) for providing TyrRS construct and Prof. Steven Rodermel (USA) for providing immutans and PTOX construct.

Funding

The initial three years of my doctorate were funded by the Erasmus Mundus BRAVE program of the European Union. I would like to acknowledge the BRAVE coordinator Dr. Andreas Voloudakis (AUA, Greece) and Prof. Dr. Sven Erik Behrens (MLU, Germany) for their constant support during the funding period.

BRAVE Colleagues

I appreciate the efforts of BRAVE fellows in Halle for making my personal life stress free and enjoyable during the stay. Thanks to Mr. Shiv Meena and Mrs. Priyanka Mishra for being great fellows, friends, and motivator since the beginning of my doctoral studies. Also thanks to Mr. Tarun K Mahto and Heena Yadav Mr. Prodyut Mondal, Mr. Tamil and Mr. Somnath Koley for being nice companions during the stay in *Halle an der Saale*.

Family

I owe a deep sense of gratitude to my parents for their love and patience during my foreign stay. Also thanks to my sister for all her encouragement and support.

Curriculum Vitae

Mayank Sharma

Date of Birth- 8th January 1990; India

Education and Work Experience

2014-2019	**Doctor rerum naturalium (Dr. rer. nat.)** (Prof. Ralf Bernd Klösgen Lab) Institute of Biology-Plant Physiology, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany
2012-2014	**Research Associate** Maharashtra Hybrid Seed Company Ltd., Jalna (MH), India
2010-2012	**Master of Science** (Plant Molecular Biology and Biotechnology) University of Delhi South Campus, New Delhi, India
2007-2010	**Bachelor of Science** (Hons.) Botany Hansraj College, University of Delhi, India
2007	**Secondary School Certificate** CBSE, New Delhi, India

Awards/Fellowships

1. Erasmus Mundus 'BRAVE' Ph.D. fellowship by European Union
2. University Medal in M.Sc. for securing highest grade
3. CSIR-UGC NET Qualification for Lectureship, University Grant Comm., India
4. Science Meritorious Award 2009-10, University of Delhi, India
5. Summer Research Fellowship 2009, Indian Academy of Science

Publications

1. Sharma, M., Kretschmer, C., Lampe, C., Stuttmann, J., & Klosgen, R. B. (2019). Determining targeting specificity of nuclear-encoded organelle proteins with the self-assembling split fluorescent protein toolkit. **Journal of Cell Science**. 132:jcs230839.
2. Sharma, M., Bennewitz, B., & Klösgen, R. B. (2018). Rather rule than exception? How to evaluate the relevance of dual protein targeting to mitochondria and chloroplasts. **Photosynthesis Research**. 138:335-343.
3. Sharma, M., Bennewitz, B., & Klösgen, R. B. (2018) Dual or Not Dual? - Comparative Analysis of Fluorescence Microscopy-Based Approaches to Study Organelle Targeting Specificity of Nuclear-Encoded Plant Proteins. **Frontiers in Plant Science**. 9:1350.
4. Sharma, M. and Bhatt, D. (2015). The circadian clock and defence signalling in plants. **Molecular Plant Pathology**. 16:210-218.

Erklärung zur Dissertation

Name/Anschrift

Mayank Sharma

geb. 08.01.1990; Mathura (Indien)

Ehrenwörtliche Erklärung zu meiner Dissertation

mit dem Titel: „**Determining targeting specificity of nuclearly encoded organelle proteins** “

Sehr geehrte Damen und Herren,

hiermit erkläre ich, dass ich die beigefügte Dissertation selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel genutzt habe. Alle wörtlich oder inhaltlich übernommenen Stellen habe ich als solche gekennzeichnet.

Ich versichere außerdem, dass ich die beigefügte Dissertation nur in diesem und keinem anderen Promotionsverfahren eingereicht habe und, dass diesem Promotionsverfahren keine endgültig gescheiterten Promotionsverfahren vorausgegangen sind.

_____ _____

Ort, Datum Unterschrift