A molecular genetic investigation into the regulation of the chloroplast TOC complex



Samuel James Watson

St Cross College

The University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy in Plant Sciences

Trinity 2019

A molecular genetic investigation into the regulation of the chloroplast TOC complex

Samuel James Watson

St Cross College

DPhil in Plant Sciences

Trinity 2019

Abstract

Chloroplasts are ancient organelles that house photosynthesis in all green plants. Most proteins required for chloroplast function are imported into the organelles from the cytosol via the coordinate action of two translocons – the TOC and TIC complexes. A recent forward genetic screen revealed that the TOC complex is conditionally targeted for degradation via the action of SP1, a chloroplast-resident E3 ubiquitin ligase. SP1 functions within the newly-described CHLORAD pathway – a pathway for the ubiquitination, membrane extraction, and degradation of chloroplast proteins. The CHLORAD pathway is activated by stress and developmental cues, but the mechanistic basis of this regulation is poorly understood. In this thesis, a genetic and functional link between chloroplast protein import, CHLORAD, and the SUMO system is described. *Arabidopsis* mutants representing most components of the SUMO pathway are shown to suppress the pale phenotype of TOC complex mutant *ppi1* with respect to chlorophyll accumulation. This suppression is also linked to increases in the accumulation of TOC proteins and enhanced chloroplast development. Conversely, the overexpression of SUMO1 or SUMO3 accentuates the *ppi1* phenotype. Protein/protein interaction studies demonstrate that key components of the SUMO system (the E2 SUMO conjugating enzyme, two E3 SUMO ligases) physically interact with TOC proteins at the chloroplast outer envelope, and SUMOylation assays indicate that TOC proteins are directly SUMOylated. This work is complemented by a forward genetic screen for additional

regulators of the TOC complex. In this screen, a pale TOC complex mutant, *ppi2-3* is chemically mutagenized and 28 viable suppressor mutants are identified, named, and characterised. Together, these two projects provide new insights into the regulation of chloroplast protein import and, perhaps, the CHLORAD pathway.

Acknowledgements

Firstly, I am grateful to the BBSRC for funding my DPhil project. I am also grateful to the Oxford Interdisciplinary Bioscience Doctoral Training Programme for providing structured support and training over the last four years.

I would like to thank my two supervisors, Professor Paul Jarvis and Professor Langdale, for their unwavering support and guidance.

The work described in this thesis could not have been completed without the generous help and support of many people in the department, but I am particularly indebted to Ms Na Li and Dr Qihua Ling. The technical help both have provided has been invaluable.

I would also like to thank my family for their support and encouragement over the last four years, and I dedicate this thesis to my wonderful grandfather, Angus Watson, who sadly passed away while I was writing it.

Abbreviations

(v/v)	Volume/Volume
(w/v)	Weight/Volume
5’ UTR	5’ untranslated region
APS	Ammonium persulfate
BiFC	Bimolecular fluorescence complementation
BN-PAGE	Blue native polyacrylamide gel electrophoresis
Bp	Base pairs
CAB	Chlorophyll a binding protein
CDS	Coding sequence
CHLORAD	Chloroplast associated protein degradation
Col-0	Columbia-0
Cia5	Chloroplast protein import apparatus 5
Clp	Caseinolytic protease
dH_2O	Distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleoside triphosphates
DTT	Dithiothreitol
DUB	Deubiquitinating enzymes
EDTA	Ethylenediaminetetraacetic acid

EGT	Endosymbiotic gene transfer
eIF4E1	Eukaryotic translation initiation factor 4E-1
EMS	Ethyl methanesulfonate
ER	Endoplasmic reticulum
ERAD	Endoplasmic reticulum associated protein degradation
fts1	Ferrochelatase two suppressor 1
FX	Filial generation x
G	G-force
GA	Gibberellic acid
gDNA	Genomic DNA
HA	Human influenza hemagglutinin
HECT	Homologous to E6AP C-terminus
HIP	Hsp70 interacting protein
HMGCR	3-hydroxy-3-methyl-glutaryl-coenzyme A reductase
HOP	Hsp70/90 organising protein
Hsp	Heat shock protein
Hz	Hertz
IP	Immunoprecipitate
kDa	Kilodalton
KOC1	Kinase at the outer chloroplast membrane 1
LB	Lysogeny broth

MAD	Mitochondrion associated protein degradation
MES	2-(N-morpholino)ethanesulfonic acid
MS	Murashige and Skoog media
Mx	Plant generations after mutagenesis
OEP64	Outer envelope protein of 64 kDa
Oep7	Outer envelope membrane protein 7
Omp85	Outer membrane protein of 85 kDa
OTS1/2	Overly tolerant to salt 1/2
OX	Overexpression
PCR	Polymerase chain reaction
PEG 4000	Polyethylene glycol 4000
PFA	Paraformaldehyde
PIC	Protease inhibitor cocktail
POTRA	Polypeptide transport associated
ppi1	Plastid protein import 1
ppi2	Plastid protein import 2
ppi3	Plastid protein import 3
Rad52	Radiation sensitive 52
RING	Really interesting new gene
RNA	Ribonucleic acid
RNF4	Ring finger protein 4

ROS	Reactive oxygen species
RT-PCR	Reverse transcriptase polymerase chain reaction
RuBisCo	Ribulose-1,5-bisphosphate carboxylase/oxygenase
SAE	SUMO activating enzyme
SCB	Sodium cacodylate buffer
SCE1	SUMO conjugating enzyme 1
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SIM	SUMO-interacting motif
SNP	Single nucleotide polymorphism
sop1-28	Suppressor of *ppi2-3* 1 to 28
SP1	Suppressor of *ppi1* 1
SP2	Suppressor of *ppi1* 2
SPF1/2	SUMO proteases related to fertility 1/2
SP-RING	SIZ1/PIAS RING domain
STUbL	SUMO-targeted E3 ubiquitin ligase
SUMO	Small ubiquitin-like modifier protein
TBE buffer	Tris/Borate/EDTA buffer
TBS	Tris-buffered saline
TBS-T	Tris-buffered saline with Tween 20
T-DNA	Transfer DNA

TE	Tris/EDTA buffer
TEM	Transmission electron microscopy
TGS	Tris/Glycine/SDS
TIC	Translocon on the inner chloroplast membrane
TOC	Translocon on the outer chloroplast membrane
TPR	Tetratricopeptide repeat
ULP	Ub-like protease
UPS	Ubiquitin-proteasome system
YFP	Yellow fluorescent protein

Table of Contents

1 General introduction 14
1.1 Plastids – evolutionary origins, functions, and diversity 14
1.1.1 What are plastids? 14
1.1.2 Plastids evolved from cyanobacteria 14
1.1.3 Plastids are diverse 15
1.2 Plastid protein import 15
1.2.1 The TOC complex 16
1.2.2 The TIC apparatus 23
1.3 The regulation of the TOC complex by the ubiquitin proteasome system 30
1.3.1 The Ubiquitin-Proteasome System (UPS) 30
1.3.2 SP1 33
1.3.3 SP2 34
1.3.4 ERAD 36
1.3.5 CHLORAD 37
1.3.6 How is the CHLORAD pathway regulated? 38
1.4 The SUMO system 38
1.4.1 The SUMO system in plants 39
1.4.2 The SUMO system, Cdc48 and ubiquitin 45
1.5 Thesis aims and structure 46

2 Genetic experiments reveal a link between plastid protein import and the SUMO system 48
2.1 Abstract 48
2.2 Introduction 49
2.3 Results 51
2.3.1 The *sce1-4* mutation suppresses the *ppi1* phenotype in a specific manner 51
2.3.2 The *siz1-4* mutant suppresses the *ppi1* phenotype in a specific manner 52
2.3.3 The *sum1-1* and *sum2-1* mutants additively suppress the *ppi1* phenotype 54
2.3.4 SUMO1 overexpression enhances the *ppi1* phenotype 55
2.3.5 The *sum3-1* mutant cannot suppress the *ppi1* phenotype 55
2.3.6 SUMO3 overexpression enhances the *ppi1* phenotype 56

2.4 Discussion 57

2.5 Figures and tables 61

3 Molecular evidence that TOC proteins physically interact with the SUMO system and are SUMOylated 78

3.1 Abstract 78

3.2 Introduction 79

3.3 Results 80

3.3.1. The E2 SUMO conjugating enzyme SCE1 physically interacts with TOC proteins in a BiFC experiment 80

3.3.2. The E2 SUMO conjugating enzyme SCE1 physically interacts with TOC proteins in an immunoprecipitation experiment 81

3.3.3 The E3 SUMO ligase SIZ1 physically interacts with TOC proteins in a BiFC experiment 82

3.3.4 A construct encoding the E3 SUMO ligase SIZ1 failed to express in an immunoprecipitation experiment 82

3.3.5 The E3 SUMO ligase HPY2 physically interacts with TOC proteins in a BiFC experiment 83

3.3.6 The E3 SUMO ligase HPY2 physically interacts with Toc159 in an immunoprecipitation experiment 84

3.3.7 Toc159 physically interacts with SUMO proteins *in vivo* 84

3.3.8 Toc33 physically interacts with SUMO proteins *in vivo* 85

3.4 Discussion 87

3.5 Figures and tables 93

4 A forward genetic screen for new regulators of the stability and function of the TOC complex 103

4.1 Abstract 103

4.2 Introduction 104

4.3 Results 107

4.3.1 Phenotypic characterisation of *ppi2-3* 107

4.3.2 Identification of new *ppi2-3* suppressor mutants 107

4.3.3 Characterisation of the new *ppi2-3* suppressor mutants 108

4.4 Discussion 111

4.4.1 Future directions 112

4.4.2 Possible suppressor loci 113
4.5 Figures and tables 115

5 General discussion 129
5.1 Summary of key findings 129

5.2 The relationship between SUMO, plastid protein import, and CHLORAD 131

5.2.1 Does SUMO recruit Cdc48 to the TOC complex? 132
5.2.2 Does SP1 act as a STUbL? 133
5.2.3 Is SP1 SUMOylated? 134
5.3 New suppressors of Toc159 mutant *ppi2-3* 135

5.4 Concluding remarks 136

6 Materials and methods 137
6.1 Plant growth conditions and physiological studies 137

6.1.1 Plants grown on media 137
6.1.2 Plants grown on soil 138
6.1.3 Crossing plants 138
6.1.4 Chlorophyll measurements 138
6.2 Plant DNA and RNA work 139

6.2.1 DNA extraction from plant tissue 139
6.2.2 PCR and PCR genotyping 139
6.2.3 Agarose gel electrophoresis 140
6.2.4 RNA extraction from plant tissue 141
6.2.5 First-strand cDNA synthesis 141
6.2.6 Semi-quantitative RT-PCR 141
6.3 Bacterial work and molecular cloning 142

6.3.1 *E. coli* strains, growth conditions, and DNA isolation 142
6.3.2 *Agrobacterium* growth conditions and transformation 143
6.3.3 Restriction enzyme-based cloning 144
6.3.4 Gateway cloning 146
6.4 Protein analysis 147

6.4.1 Protein extraction 147
6.4.2 SDS-PAGE 148
6.4.3 Coomassie staining 149

6.4.4 Western blotting 149
6.5 Transient expression assays and interaction studies 150

6.5.1 Protoplast isolation 150
6.5.2 Protoplast transfection 151
6.5.3 BiFC experiments 152
6.5.4 Immunoprecipitation 153
6.6 Transmission electron microscopy 154

6.6.1 Sample fixation 154
6.6.2 Sample dehydration 155
6.6.3 Epoxy resin infiltration 156
6.6.4 Sample embedding and sectioning 156
6.6.5 Imaging 157
6.7 The mutant screen 157

6.7.1 EMS mutagenesis 157
6.7.2 M2 screening 158
6.8 Tables 159

7 Bibliography 163

1 General introduction

1.1 Plastids – evolutionary origins, functions, and diversity

1.1.1 What are plastids?

Plastids are a family of organelles that are exclusively found in the cells of plants, algae, and apicomplexan parasites (Jarvis and Lopez-Juez, 2013). The archetypal plastid is undoubtedly the chloroplast, which houses photosynthesis in all autotrophic eukaryotes. Other well-studied plastid types include the amyloplast, which stores starch and is therefore of critical importance to crop productivity, and the chromoplast, which accumulates the carotenoid pigments that give fruit and some flowers their characteristic colours.

1.1.2 Plastids evolved from cyanobacteria

Plastids have an unusual evolutionary history: they are the integrated descendants of a free-living ancestor that was related to extant cyanobacterial species. The ancient plastid ancestor entered the eukaryotic lineage via endosymbiosis more than one billion years ago (Yoon et al., 2004). The precise timing of the initial endosymbiotic event is contentious. Fossils claimed to represent eukaryotic algae can be found in rocks that are thought to be more than two billion years old (Han and Runnegar, 1992). Relatives of the extant red algal genus *Bangia* can be found in rocks that are up to 1.25 billion years old (Butterfield et al., 1990). Whatever the exact timing of the event, its importance to the development of complex autotrophic life cannot be overstated. Aside from one arguable exception (Marin et al., 2005), the endosymbiotic event gave rise to all modern plants and algae.

Vestiges of the plastid's evolutionary history can be found in many of the structural and molecular features common to all plastid types: for example, a miniaturised genome, a double membrane, and a number of proteases of bacterial origin (Jarvis and Lopez-Juez, 2013; Nishimura et al., 2016). The plastid genome is still functional, and it encodes

approximately 100 proteins. However, the genes encoding almost all the proteins required for plastid function are now housed by the central, nuclear genome. These genes were transferred to the nuclear genome via endosymbiotic gene transfer (EGT).

1.1.3 Plastids are diverse

In higher plants, plastids are a structurally and functionally diverse group of organelles (for a detailed review, see: (Lopez-Juez and Pyke, 2005; Jarvis and Lopez-Juez, 2013)). All plastids develop from the etioplast, a small precursor organelle that can differentiate into various mature plastid types in response to environmental and developmental cues. The proplastid can develop into the amyloplast, a plastid specialised for starch storage; the elaioplast, an organelle specialised for the synthesis and storage of important lipids; and the chloroplast, an organelle that houses photosynthesis as well as the synthesis of a large number of indispensable molecules (for example, amino acids). In addition, the proplastids of plants grown in the dark can differentiate into etioplasts. Etioplasts accumulate chlorophyll precursors and rapidly differentiate into chloroplasts when exposed to light (Pogson et al., 2015).

The chloroplast can, itself, differentiate into different plastid types. During senescence, chloroplasts differentiate into gerontoplasts. Gerontoplasts break down the photosynthetic machinery, allowing the constituent molecules to be recycled. Chloroplasts can also differentiate into chromoplasts, organelles that synthesise and store pigments and give fruit and some flowers their colours (Sadali et al., 2019).

1.2 Plastid protein import

As the majority of the plastid genome is now housed by the central, nuclear genome, genes required for plastid function must be transcribed in the nucleus, translated in the cytosol, and their products imported into plastids in a post-translational manner. This import necessitates a machinery that can facilitate the movement of proteins across the plastid double envelope membrane system. It is now known that this machinery

comprises two large, proteinaceous complexes that reside in the chloroplast outer and inner envelope membranes. These complexes are known as the translocon on the outer chloroplast membrane (TOC) and the translocon on the inner chloroplast membrane (TIC) complexes (Jarvis, 2008).

The TOC and TIC complexes act cooperatively to engage and import plastid-targeted preproteins. Each preprotein is synthesised with an N-terminal 'transit peptide', a small amino acid sequence that is recognised by the plastid import machinery (Jarvis, 2008). After import, transit peptides are removed via the action of a dedicated protease.

The TOC and TIC complexes are of fundamental significance to this project, and they will be discussed in detail next.

1.2.1 The TOC complex

The TOC complex resides in the chloroplast outer envelope membrane and mediates the translocation of preproteins across the membrane. The TOC complex comprises three major components, named in accordance with their predicted molecular weights: Toc159, Toc34 and Toc75 (Jarvis, 2008). The three components have distinct roles within the plastid import process. Toc159 and Toc34 are thought to function as preprotein receptors, and both contain large globular GTPase domains that extend into the cytosol (Jarvis, 2008). Toc75 forms a membrane-spanning channel, and contains a large, membrane-spanning β-barrel domain and three polypeptide-transport associated (POTRA) domains that reside in the intermembrane space and are likely to act as chaperones for membrane-transiting preproteins (O'Neil et al., 2017). The molecular weight and component stoichiometry of the complex is still debated. Estimations of the size of the complex show considerable variation: ~500 kDa in one study, and 800-1000 kDa in another (Schleiff et al., 2003; Kikuchi et al., 2006). The same studies proposed similarly different stoichiometry estimates: a 4:4:1 vs a 6:6:2 Toc34:Toc75:Toc159 arrangement.

The sequencing of the genome of the model plant *Arabidopsis thaliana* enabled the identification of genes encoding TOC complex components (Jarvis et al., 1998; Jackson-Constan and Keegstra, 2001; Baldwin et al., 2005). Intriguingly, the three core TOC proteins are encoded by small gene families that display a degree of functional diversification, as revealed through genetic and biochemical studies.

1.2.1.1 Toc159

Toc159 is a large and complex preprotein receptor. It contains three major domains: A C-terminal membrane anchoring domain (known as the 'M' domain), a GTPase domain (the 'G' domain), and an N-terminal, highly acidic domain of poorly understood function (the 'A' domain) (Bolter et al., 1998; Chen et al., 2000; Andres et al., 2010).

The A and G domains extend into the cytosol. The A domain is readily removed from the protein via the action of an unknown protease; this occurs so frequently that Toc159 was initially described as an 86 kDa protein (the molecular weight of the protein without the A-domain) (Perry and Keegstra, 1994). Although the function of the A-domain is still incompletely understood, domain-swapping experiments with the closely related Toc132 protein (described later) imply a role in selective preprotein recognition (Inoue et al., 2010). As will be discussed later, TOC complexes composed of different components show distinct client specificities, with some variants specialised in the import of highly abundant proteins involved in photosynthesis, and others specialised in the import of less abundant, housekeeping proteins. The A-domain may have a role in this process, as the client preferences of Toc159 and Toc132 can be reversed by swapping their respective A-domains. In addition, although the A-domain is nonessential (Schnell et al., 1994), one study found that full-length Toc159 was more efficient at facilitating the import of preproteins than the proteolytically cleaved variant (Bolter et al., 1998).

The presence of a GTPase domain implies that Toc159's activity is regulated by a GTPase cycle. Evidence for this mode of regulation was provided by a complementation

study (Agne et al., 2009); however, the results of this work were less striking than might be expected. The authors were able to demonstrate that the pale, albino phenotype of the Toc159 null mutant *plastid protein import 2* (*ppi2)* could be fully rescued via the transgenic expression of a Toc159 variant bearing a mutation in a conserved GTPase domain residue; however, the plants did show a decrease in the efficiency of chloroplast protein import in an *in vitro* experiment.

The M-domain is an atypical membrane-anchoring domain. It does not contain an obvious sequence of hydrophobic amino acids as might be expected; however, it is essential for the membrane insertion of Toc159 (Hirsch et al., 1994; Kessler et al., 1994) and for the import of proteins across the chloroplast outer envelope (Lee et al., 2003). A series of protease protection and salt/alkaline extraction experiments demonstrated that Toc159 is tightly embedded within the chloroplast outer membrane (Hirsch et al., 1994; Kessler et al., 1994). In addition, a proteoliposome reconstitution study found that Toc75 was essential for the membrane integration of Toc159, leading to the proposal that the protein is inserted within the TOC complex (i.e. proximate to or within the Toc75 pore) (Wallas et al., 2003).

1.2.1.2 The Toc159 gene family

Toc159 is the best-studied member of a small gene family. Four genes encode Toc159 family proteins in *Arabidopsis*: Toc159, Toc132, Toc120 and Toc90 (Kubis et al., 2004). Like other TOC proteins, these were named in accordance with their predicted molecular weights. Null mutants of the genes encoding all four proteins have been identified and described, and this has provided insight into each gene's function.

Toc159 can be considered the 'dominant' member of the family. Its function is indispensable; the Toc159 null mutant *ppi2* is albino and cannot survive to adulthood (Bauer et al., 2000). Chloroplast biogenesis is completely abrogated by the *ppi2* mutation. Toc159 is strongly expressed in young, photosynthetic tissue, and the protein shows a

receptor preference for highly abundant photosynthetic preproteins (Bauer et al., 2000; Kubis et al., 2004).

Toc132 and Toc120 are highly similar proteins. A phylogenetic analysis divided the Toc159 family into two subfamilies, one containing Toc159 and the other Toc132/Toc120 (Kubis et al., 2004). Unlike Toc159, Toc132 and Toc120 show relatively uniform patterns of expression with no particular bias towards young, photosynthetic tissues. In most tissues, Toc132 displays a lower level of expression than Toc159, and Toc120 in turn displays a lower level of expression than Toc132. Toc132 null mutants display a mildly chlorotic, reticulate phenotype that is only apparent in mature plants. Toc120 null mutants show no obvious aberrant phenotype. However, double Toc132/Toc120 null mutants are non-viable, indicating that the two genes are collectively essential and display functional redundancy (Ivanova et al., 2004; Kubis et al., 2004).

Toc90 is the smallest member of the Toc159 family, and it is structurally distinctive. While every other protein in the family contains an A-domain, this is completely lacking in Toc90 (Kubis et al., 2004). Despite the truncated nature of Toc90, it is expressed at a uniformly high level throughout a range of tissue types. A null mutant, *ppi4*, has been identified and described (Hiltbrunner et al., 2004). This mutant has no obvious phenotypic defects; however, it can enhance the severity of the albino phenotype of the Toc159 mutant *ppi2*. In addition, the *ppi2 ppi4* double mutant shows an even greater decrease in the accumulation of chlorophyll a/b binding protein (CAB), a key photosynthetic protein. This work was complemented by another study that demonstrated that the overexpression of Toc90 can partially rescue the phenotype of *ppi2* with respect to chlorophyll concentration and the import of photosynthetic preproteins (Infanger et al., 2011). Taken together, these studies indicate that Toc90 shows some degree of functional redundancy with Toc159, but that it is nonessential for plant viability under standard growth conditions.

1.2.1.3 Toc34 and Toc33

Toc34 is a receptor GTPase that associates with the other TOC complex components. Like Toc159, Toc34 was first identified in biochemical experiments carried out on pea chloroplasts and it is the founding member of a small family of similar proteins (Kessler et al., 1994). Two such proteins have been described in *Arabidopsis*: Toc34 and Toc33. Confusingly, Toc33 is the 'dominant' isoform in *Arabidopsis*.

Toc34 is a relatively small protein with a comparatively simple domain structure. Toc34 contains two major domains: a C-terminal transmembrane anchoring domain, and an N-terminal GTPase domain that extends into the cytosol (Kessler et al., 1994; Constan et al., 2004). A crystal structure of the soluble domain of Toc34 confirmed the GTP-binding activity of the protein and demonstrated that it can dimerise, at least *in vitro* (Sun et al., 2002). Toc34 is weakly expressed in a range of plant tissues, but shows a high rate of expression in roots (Gutensohn et al., 2000; Kubis et al., 2003). Two Toc34 null mutants have been described - *ppi3-1* and *ppi3-2* – but surprisingly, neither of them displays any obvious phenotypic defects in aerial tissues. Similarly, chlorophyll size, ultrastructure, and protein import are not altered in either mutant. However, both mutations induce a subtle root growth defect, which is consistent with the expression pattern of Toc34. The lack of an obvious mutant phenotype in the aerial tissues of *ppi3* mutants indicates that the closely related protein Toc33 can compensate for the lack of Toc34 in these parts of the plant.

Toc33 was first identified in a forward genetic screen for mutants defective in plastid biogenesis (Jarvis et al., 1998). In this study, a Toc33 null mutant – *plastid protein import 1* (*ppi1*) – was described. The *ppi1* mutant is small, chlorotic, and displays a specific defect in the import of photosynthetic preproteins into the chloroplast (Kubis et al., 2003). Toc33 has a similar domain structure to Toc34, but displays a different expression profile (Gutensohn et al., 2000; Kubis et al., 2003). Toc33 shows a higher general level of expression than Toc34, and is expressed particularly strongly in young, photosynthetic

tissue. As discussed, Toc33 and Toc34 show a degree of functional redundancy. The double null mutant *ppi1 ppi3* is non-viable, demonstrating that the two genes are collectively essential for plant viability (Constan et al., 2004).

1.2.1.4 Toc75

Toc75 is a central component of the TOC complex, and it is a member of the outer membrane protein of 85 kDa (Omp85) superfamily of β-barrel membrane proteins (Schnell et al., 1994; Tranel et al., 1995; O'Neil et al., 2017) . Omp85 family proteins are found in the outer membranes of Gram-negative bacteria, as well as the outer membranes of chloroplasts and mitochondria (a reflection of the organelles' endosymbiotic origin) (Gentle et al., 2005). Toc75 functions as a membrane channel, physically permitting the transport of preproteins across the chloroplast outer envelope membrane (Hinnah et al., 2002). The domain architecture of Toc75 comprises a C-terminal, membrane-embedded β-barrel domain that is predicted to contain 16 transmembrane β-strands, and an N-terminal region that localises to the intermembrane space and contains three POTRA domains (O'Neil et al., 2017). The function of the β-barrel domain of Toc75 is uncontroversial: it forms the membrane-spanning pore of the TOC complex. The function(s) of the three POTRA domains took longer to resolve. The three domains are essential for plant viability, as demonstrated in a series of domain deletion experiments (Paila et al., 2016). A series of constructs encoding truncated variants of Toc75 lacking 1-3 POTRA domains failed to complement a non-viable Toc75 null mutant. The same study found that the domains promote preprotein import in an *in vitro* import assay. The crystal structure of the three POTRA domains has been solved, and each retains a highly conserved β1α1α2β2β3 fold that is typical of POTRA family proteins (O'Neil et al., 2017). This work also demonstrated that the POTRA domains can physically interact with preproteins, and that POTRA2 and POTRA3 collectively display chaperone-like activity in an insulin aggregation assay.

As is the case with the other major TOC proteins, multiple genes encode Toc75 family proteins in *Arabidopsis*. Three genes have been identified and described: Toc75-I, Toc75-III, and Toc75-IV (Baldwin et al., 2005). Toc75-III is the 'dominant' isoform, and it is strongly expressed throughout the plant with a particular bias towards young, rapidly growing photosynthetic tissues. Toc75-IV is weakly expressed, and it shows a more uniform expression profile than Toc75-III. Toc75-I is almost certainly a pseudogene – its sequence contains a series of nonsense, frameshift, and splice-junction mutations and it is not obviously expressed in any tissue. Toc75-III and Toc75-IV mutants have been identified and characterised. Null mutants of Toc75-III are non-viable in the homozygous state, reflecting its abundant expression throughout the plant. Interestingly, a weak allele was identified in a forward genetic screen for regulators of root gravitropism (Stanga et al., 2009). Although the relationship between gravity-sensing and the TOC complex is still poorly understood, this mutant – known as *toc75-III-3* – has been a useful resource for the chloroplast protein import field. As is the case with other major TOC mutants, *toc75-III-3* is chlorotic due to a defect in the import of chloroplast-targeted preproteins (Huang et al., 2011). In addition, the mutation causes defects in chloroplast biogenesis and the biogenesis of other TOC proteins. Conversely, Toc75-IV null mutants have no obvious aberrant phenotype, and a series of genetic experiments demonstrated that null mutants do not modify the phenotypes of several other tested TOC mutants (Baldwin et al., 2005). However, the etioplasts of Toc75-IV null mutants show subtle structural changes and the plants display de-etiolation defects. These findings may indicate a niche role for the protein within certain developmental processes.

1.2.1.5 Toc64 – a functional component of the TOC complex?

The Toc64 protein resides in the chloroplast outer envelope membrane and has been proposed to function as an initial receptor for chloroplast-targeted preproteins that are complexed with the cytosolic chaperone Hsp90 (Sohrt and Soll, 2000; Qbadou et al., 2006). Hsp90 chaperones preproteins in the cytosol alongside at least two other proteins,

Hsp70 and 14-3-3 (Flores-Perez and Jarvis, 2013). However, a study found that Toc64 null mutations do not cause any obvious phenotypic defects in *Arabidopsis* (Aronsson et al., 2007). In this study, mutants were carefully described with respect to various parameters including size, chlorophyll content, and plastid protein import efficiency. In all cases, the mutants were indistinguishable from wild-type controls. In addition, a second study explored the consequences of knocking out Toc64 in the lower plant *Physcomitrella patens* (Hofmann and Theg, 2005). Although the authors noticed a minor change in chloroplast shape in the Toc64 null mutants, the plants were otherwise indistinguishable from wild type, and showed no obvious change in chlorophyll content or chloroplast protein import. Taken together, these studies call into question the role of Toc64 – or, as it is now sometimes known outer envelope protein 64 kDa (OEP64) – within the plastid import process.

1.2.2 The TIC apparatus

While the core TOC proteins have been identified and the basic composition of the complex is uncontroversial, the TIC apparatus is still incompletely understood. In addition, major controversies exist within the field and no consensus model of TIC composition and function has yet emerged. Many putative TIC proteins have been identified and studied; however, even basic questions (such as the nature of the TIC pore) remain contentious. Proteins that have been associated with the TIC complex include: Tic110, Tic40, Tic22, Tic21, Tic20, Tic62, Tic55, Tic32 (reviewed by Jarvis (2008)) and, more recently, Tic214, Tic100, Tic56 (Kikuchi et al., 2013), and Tic236 (Chen et al., 2018). Due to the controversies within the field, this section will be structured in a different way to the preceding section. First, the ‘classical’ model and its components will be described, before the new model proposed by Professor Masato Nakai’s group will be explored. It is, however, important to note two things: first, that there was not a clear consensus model of TIC function before the publication of the Nakai model, and second, that the two models are not necessarily mutually exclusive.

1.2.2.1 The TIC field prior to the 1-megadalton complex controversy

1.2.2.1.1 Tic110

Tic110 was the first described TIC protein, and it was identified in a series of biochemical experiments using isolated pea chloroplasts (Schnell et al., 1994; Kessler and Blobel, 1996). Tic110 localises to the inner chloroplast envelope and, while its topology has been disputed, it is broadly accepted that it contains two membrane-spanning α-helices and a large, hydrophilic stromal extension (Jackson et al., 1998; Inaba et al., 2003). This stromal domain can bind transit peptides and may act to recruit chaperones to trafficking preproteins (Kessler and Blobel, 1996; Akita et al., 1997; Nielsen et al., 1997; Inaba et al., 2003). More controversially, it has been proposed that Tic110 may form at least part of the membrane-spanning pore of the TIC complex. This proposal is based on an electrophysiological study carried out using liposome-reconstituted recombinant proteins (Heins et al., 2002). However, a subsequent study resolved the crystal structure of various fragments of Tic110 and concluded that the protein's structure is inconsistent with the proposed channel function (Tsai et al., 2013). Instead, this study found that the stromal domain of Tic110 was structurally similar to proteins that have been proposed to act as scaffolds for protein-protein interactions. Interestingly, Tic110 physically interacts with the heat shock protein 100 (Hsp100)-type chaperone Hsp93/ClpC, and this may be been linked to a preprotein quality control function at the stromal side of the chloroplast inner membrane (Flores-Perez et al., 2016).

1.2.2.1.2 Tic22

Tic22 is a well-established component of the plastid import machinery. The small protein was first identified in a series of cross-linking experiments; these studies found that Tic22 could physically interact with transiting preproteins (Kouranov and Schnell, 1997; Kouranov et al., 1998). Tic22 resides in the intermembrane space and shows a peripheral association with the chloroplast outer envelope; it is therefore likely to be one

of the first TIC proteins encountered by trafficking preproteins and may act to coordinate the function of the TOC and TIC complexes.

Tic22 family proteins are encoded by two genes in *Arabidopsis*: Tic22-III and Tic22-IV (Jackson-Constan and Keegstra, 2001). Two complementary studies demonstrated that the two proteins are functionally redundant and important for chloroplast protein import (Kasmati et al., 2013; Rudolf et al., 2013). Null mutants of both genes were obtained and described in both cases. Surprisingly, the two studies are not in complete agreement as to the phenotypes of the mutants. This may be due to the different T-DNA insertion mutants used in the two studies. In the first, no mutant phenotype was observed in the single Tic22-III null mutant, whereas in the second the plants were reported to display a weakly chlorotic phenotype. In addition, while the second study found that the double Tic22-III/IV null mutant showed severe chlorosis, the first found that the chlorotic phenotype was only apparent under high light stress. Nonetheless, both studies demonstrated that both Tic22 isoforms are strongly expressed throughout the plant, are functionally redundant, and show defects in chloroplast protein import, firmly establishing the protein as an important component of the chloroplast protein import machinery.

1.2.2.1.3 Tic20

Tic20 was discovered in parallel with Tic22 (Kouranov and Schnell, 1997). This work found that preproteins are physically close to Tic20 during transit. Tic20 is a small protein, and it contains four transmembrane α-helices (Chen et al., 2002). These structural features have led several groups to propose that Tic20 is the major TIC complex channel protein (Kouranov et al., 1998; Kikuchi et al., 2013). This view was supported by a proteoliposome reconstitution study which found that Tic20 can form a cation-selective channel (Kovacs-Bogdan et al., 2011). In addition, a biochemical study that found that Tic20 is present in a large complex that physically associates with

preproteins (Kikuchi et al., 2006). This study relates to the major controversy within the TIC field that will be discussed in detail later.

Tic20 is encoded by four genes in *Arabidopsis*: *TIC20-I*, *TIC20-II*, *TIC20-IV* and *TIC20-V* (Kasmati et al., 2011). All four genes encode proteins that localise to the chloroplast inner membrane. However, only one, *TIC20-I*, induces a clear aberrant phenotype when mutated (Hirabayashi et al., 2011; Kasmati et al., 2011). The *tic20-I* mutation induces a severe albino phenotype which is accompanied by obvious defects in chloroplast biogenesis. In addition, a group using an antisense approach found that a Tic20-I knockdown line showed retardation of chloroplast preprotein import (Chen et al., 2002). As discussed, null mutations in any of the other Tic20 genes do not cause obvious mutant phenotypes, at least under standard growth conditions. However, double *tic20-I tic20-IV* null mutants are non-viable, indicating that Tic20-IV may show partial functional redundancy with Tic20-I. In support of this view, Tic20-IV overexpression can partially complement the *tic20-I* null mutant; although the resulting plants do not show any obvious change in chlorophyll concentration, they grow considerably larger than control *tic20-I* plants (Kasmati et al., 2011).

1.2.2.1.4 Tic21

Tic21 is a small protein that resides in the chloroplast inner envelope and physically associates with TOC and TIC proteins (Teng et al., 2006). Although Tic21 is of a similar size to Tic20, the two proteins are not closely related. However, Tic21 has, like Tic20, four putative transmembrane spans, and so in broad structural terms, the two proteins are similar. Tic21 was identified in a forward genetic screen for mutants defective in chloroplast protein import, and, correspondingly, the Tic21 null mutant *chloroplast protein import apparatus 5* (*cia5*) displays a defect in chloroplast protein import as well as an albino phenotype. However, the function of Tic21 is disputed; one study argued that Tic21 is not involved in plastid protein import but instead has a role within the maintenance of iron homeostasis (Duy et al., 2007).

Tic21 was, like Tic20, found to associate with preproteins as part of a large import complex (Kikuchi et al., 2006). However, its precise role within the chloroplast protein import process is still unknown.

1.2.2.1.5 Tic40

Tic40 was first identified in a series of biochemical experiments in pea (Wu et al., 1994). Chemical cross-linking and co-immunoprecipitation experiments demonstrated that Tic40 co-localises with trafficking preproteins at the chloroplast inner envelope. Further studies demonstrated that Tic40 forms a complex with Tic110 and Hsp93, another protein thought to function within the chloroplast preprotein import process (Stahl et al., 1999; Chou et al., 2003). Tic40 has a single N-terminal transmembrane domain and a large stromal extension (Stahl et al., 1999; Chou et al., 2003). The stromal region contains a series of tetratricopeptide repeat (TPR) domains; TPR domains are known to promote protein-protein interactions. The stromal region also contains a region with homology to the mammalian co-chaperones Hsp70-interacting protein (Hip) and Hsp70/Hsp90-organising protein (Hop) (Frydman and Hohfeld, 1997). The presence of these structural features imply that Tic40 might act as a scaffold for the recruitment of chaperones during preprotein import (Bedard et al., 2007).

Tic40 is encoded by a single gene in *Arabidopsis* (Chou et al., 2003; Kovacheva et al., 2005). Null mutants have been described, and these show moderately severe chlorosis and dwarfism; however, the plants are viable and fertile. In addition, the mutants show defects in chloroplast structure and preprotein import, strongly indicating an important – but not entirely essential - role for the protein within chloroplast preprotein import.

1.2.2.1.6 Tic236

Tic236 is a new component of the chloroplast import machinery, and it was recently discovered in a series of elegant biochemical experiments (Chen et al., 2018). Tic236 is a large and highly conserved protein that is anchored in the chloroplast inner

membrane and extends into the intermembrane space. The large, intermembrane space domain of Tic236 physically associates with the POTRA domains of Toc75, while the rest of the protein physically associates with the TIC complex. The protein thus acts as a bridge, linking the two chloroplast translocons. Tic236 null mutants are non-viable, and hypomorphic mutants are chlorotic and show defects in chloroplast protein import, indicating that the 'bridging' of the TOC and TIC complexes is functionally important.

1.2.2.1.7 Tic62, Tic55 and Tic32 – TIC-associated redox regulators?

Tic62, Tic55, and Tic32 are three proteins that were proposed to physically interact with Tic110 (Caliebe et al., 1997; Hormann et al., 2004; Stengel et al., 2008). The proteins contain domains that are associated with redox-related processes (for example, Tic55 contains a Rieske-type Fe–S cluster binding site), and, in some cases, the functionality of these domains has been demonstrated. It has been argued that the three proteins function in the redox regulation of chloroplast protein import (Balsera et al., 2010); however, this is not broadly accepted.

1.2.2.2 The 1-megadalton complex and TIC

At time of writing, the TIC field is in the midst of a major controversy. Until recently, the core components of the TIC complex – Tic110, Tic20, and Tic40 – were widely accepted, even if the individual functions of the proteins were the subject of debate. However, in 2013 a study was published that claimed to overturn much of what was known about the basic composition of the complex (Kikuchi et al., 2013). In the study, plants were transformed with a protein A-tagged variant of Tic20-I. The tagged Tic20-I was immunoprecipitated and the captured complexes were resolved via blue native polyacrylamide gel electrophoresis (BN-PAGE). Tic20-I was found to associate with three novel proteins, which were named Tic214, Tic100, and Tic56. The four proteins formed a complex that was about 1-megadalton in size. Tic100 and Tic56 null mutants were obtained, and both sets of mutants displayed albino phenotypes and an inability to accumulate photosynthetic proteins. No Tic214 mutant could be obtained, as, remarkably,

Tic214 seems to be encoded by *ycf1*, a gene housed within the chloroplast genome. Biochemical experiments suggested that all three new TIC proteins physically associate with preproteins and may form a complex when reconstituted in proteoliposomes.

The most controversial claims of the study related to Tic110 and Tic40. None of the four 1-megadalton complex proteins physically interacted with Tic110 in an immunoprecipitation experiment. In addition, a further experiment demonstrated that two model preproteins physically interact with all four 1-megadalton complex proteins, but not Tic110 or Tic40. On the basis of these observations, the group responsible for the study argued that Tic110 and Tic40 are not core components of the TIC complex, and that the basic model of the TIC complex would need to be extensively revised.

The claim that the newly described 1-megadalton complex was the bona fide TIC complex was challenged by a number of groups. One major point of contention related to the evolutionary conservation of the new TIC proteins. For example, Tic214 is an essential protein in *Arabidopsis*, but it appears to be absent in grasses (including wheat, maize, and rice) (de Vries et al., 2015). In addition, mutants of most of the previously established TIC components display defects in chloroplast protein import. A detailed examination of this controversy is beyond the scope of this review; interested readers are invited to consult the following papers (de Vries et al., 2015; Nakai, 2015a; Nakai, 2015b).

When a consensus model eventually emerges, it seems likely that the 1-megadalton complex will be reconciled with the ‘classic’ TIC components. For example, it could be that a complex comprising Tic110 and Tic40 acts downstream of the 1-megadalton complex to chaperone incoming preproteins. Whatever the truth, it is unlikely that 25 years of work will be wholly overturned.

1.3 The regulation of the TOC complex by the ubiquitin proteasome system

As discussed, most components of the plastid import machinery were first identified in biochemical experiments using isolated pea chloroplasts and model import substrates. However, the discovery of Toc33 in the late 1990s was a watershed moment for the field in that it represented one of the first successful applications of *Arabidopsis* genetics to the study of plastid preprotein import (Jarvis et al., 1998). Throughout the following decades a number of groups attempted to search for new components or regulators of the plastid import machinery using forward genetic screens. In one of these screens, the Toc33 null mutant *ppi1* (first described in the pioneering 1998 paper) was used as a starting genotype in a second-order screen. A batch of *ppi1* seeds were chemically mutagenized and 'suppressor' mutants – i.e. plants that showed phenotypic recovery with respect to chlorosis – were identified. Four mutations at two loci were identified and mapped, and the two loci were named *suppressor of ppi1 1* (SP1) and *suppressor of ppi1 2* (SP2). SP1 was the first locus to be characterised in detail, and, intriguingly, it was shown to encode an E3 ubiquitin ligase. This finding was notable as it indicated, for the first time, that the Ubiquitin Proteasome System (the UPS) directly regulates chloroplast processes.

1.3.1 The Ubiquitin-Proteasome System (UPS)

Cells must maintain a balance between protein synthesis and protein degradation to survive and function. The UPS plays an essential role in the degradation of proteins. As the UPS is of direct relevance to much of the work described in this project, it will be briefly described next. For a more detailed overview, readers are invited to read one of the following reviews, from which much of the following summary is derived (Vierstra, 2009; Komander and Rape, 2012).

Ubiquitin is a small, 76 amino acid protein that is conserved throughout the eukaryotic kingdom. Ubiquitin is attached to target proteins via the coordinate action of a

dedicated enzyme machinery in a process known as ‘ubiquitylation’ or ‘ubiquitination’. Although it is now known that ubiquitination can have a range of consequences for the substrate, the most common fate of ubiquitinated proteins is degradation via the action of the 26S proteasome, a large and highly conserved proteolytic complex. The targeted protein is rapidly degraded by the 26S proteasome and small peptide fragments and ubiquitin moieties are released for reuse.

Ubiquitination is a dynamic process that requires a large number of enzymes and cofactors. The core components of the ubiquitination machinery include three classes of enzyme that collectively activate and conjugate ubiquitin to specific substrates. First, an E1 ‘activating’ enzyme activates the ubiquitin by forming a thioester bond with a conserved glycine residue contained within the C-terminus of the ubiquitin protein. This ‘activation’ step is ATP-dependent. Second, the E1 enzyme transfers the activated ubiquitin to an E2 ‘conjugating’ enzyme via trans-esterification (Stewart et al., 2016). Third, the E2 conjugating enzyme associates with an E3 ‘ubiquitin ligase’, which, in turn, associates with the substrate and guides the E2-dependent conjugation step (Zheng and Shabek, 2017). E3 activity thus promotes substrate-selective ubiquitination, although in a restricted number of cases E2 conjugating enzymes mediate ubiquitination in an E3-independent manner. In both cases, ubiquitin is conjugated to target proteins via an isopeptide bond between a conserved glycine residue in its C-terminal region and, typically, an exposed lysine residue in the target protein. Ubiquitin contains seven lysine residues and can itself be ubiquitinated, leading to substrates bearing multimeric ubiquitin chains. Lysine-48 and Lysine-63 are common platforms for chain-building (Akutsu et al., 2016). The various modes of poly-ubiquitination are thought to promote different substrate fates; for example, Lysine-48 modification is associated with 26S-proteasomal degradation, whereas Lysine-63 modification is associated with a range of cellular activities including endocytic trafficking and translation (Akutsu et al., 2016).

Alongside the core E1/E2/E3 enzymes, a number of additional proteins and accessory factors are required for ubiquitination to proceed. A dedicated family of proteases known as the deubiquitinating enzymes (DUBs) facilitate the proteolytic removal of ubiquitin moieties from substrates. DUBs promote the maturation of newly-synthesised ubiquitin peptides, positively regulate the stability of ubiquitinated proteins, and maintain the function of the 26s proteasome (Amerik and Hochstrasser, 2004). In addition, a specialised class of proteins known as E4 ubiquitin ligases promote ubiquitin chain formation (Koegl et al., 1999).

1.3.2 The Ubiquitin-Proteasome System in plants

Plant genomes, like the genomes of all other eukaryotic organisms, encode a functional ubiquitin-proteasome system. Intriguingly, the plant UPS is encoded by an extremely large number of genes; by some estimates up to 6% of the *Arabidopsis* transcriptome encodes UPS components (Vierstra, 2009). Most of this complexity comes from an enormous repertoire of E3 ubiquitin ligases - it is thought that more than 1400 loci could encode proteins of this category in *Arabidopsis* (Vierstra, 2009).

E3 ubiquitin ligases can be placed into four broad families. These are: The Really Interesting New Gene (RING) family, the U-Box family, and the Homologous to E6AP C-Terminus (HECT) family, and the Cullin-RING Ligase family (CRL) (Vierstra, 2009). E3 ubiquitin ligases can exist as single small proteins or as components of large, multimeric complexes. Although E3 ubiquitin ligases are abundant within the cell, they are not thought to localise to the interior of membrane-bound organelles (with the exception of the nucleus). However, a large number of E3 ubiquitin ligases reside in various biological membranes, and the UPS extends its reach into organelles via the action of a series of pathways that mediate the extraction and ubiquitination of organelle-localised substrates. These pathways are of considerable importance to this project and will be discussed later.

1.3.2 SP1

As discussed in a previous section, the *SP1* locus was first identified as a second-site modifier of the pale phenotype of Toc33 null mutant *ppi1* (Ling et al., 2012). More precisely, *sp1* mutations can 'suppress' or 'rescue' the *ppi1* phenotype with respect to chlorophyll concentration, TOC protein abundance, chloroplast biogenesis, and chloroplast protein import.

The *SP1* locus encodes a RING-type E3 ubiquitin ligase that localises to the chloroplast outer envelope. The SP1 protein is small and has a relatively simple domain architecture, comprising two transmembrane domains, and intermembrane space domain that has a role in substrate recognition, and an enzymatically active RING domain that is responsible for the protein's E3 ubiquitin ligase activity (Ling et al., 2012). The RING domain extends into the cytosol, while the two transmembrane domains firmly anchor SP1 within the chloroplast outer membrane. SP1 physically interacts with all tested TOC proteins and, under certain circumstances, catalyses their ubiquitination. As SP1 promotes the turnover of TOC proteins, *sp1* mutations therefore, in the Jarvis laboratory model, suppress the *ppi1* phenotype by allowing various TOC proteins (most strikingly Toc75) to increase in abundance. The increased abundance of these TOC proteins can partially compensate for the lack of Toc33 receptor protein in the *ppi1* background.

SP1 activity appears to be important during plastid transitions, as *sp1* mutations reduce the efficiency of the proplastid-to-chloroplast transition (chloroplast biogenesis) and also the chloroplast-to-gerontoplast transition (chloroplast senescence) (Ling et al., 2012). SP1 is thought to promote these transitions by modifying the TOC complex isoform composition of the affected plastids. As discussed, the core components of the TOC complex are encoded by small gene families, and the various isoforms of each TOC protein form complexes with distinct but overlapping substrate preferences. Plastid interconversion events are driven by major changes to the organelles' proteomes, and SP1 drives these changes by modifying the composition of TOC complex isoforms in the

outer membranes of affected plastids. These changes modify the import preferences of the plastids, thus leading to proteome changes and developmental transitions.

SP1 activity is also important in plant survival in the face of environmental challenge: *sp1* mutants are hypersensitive to salt stress, osmotic stress, and oxidative stress (Ling and Jarvis, 2015). Conversely, SP1 overexpression makes plants more resistant to all three forms of stress. These growth and survival outcomes are linked to differences in the concentration of reactive oxygen species (ROS) within the leaf tissues of the plants. ROS accumulation is induced by a range of environmental challenges and causes damaging oxidative stress within the cell (Watson et al., 2018). Interestingly, *sp1* mutants show an increased rate of ROS accumulation in response to a range of abiotic stressors, and, conversely, transgenic plants that overexpress SP1 show a reduced rate of ROS accumulation under similar circumstances. SP1 is thought to reduce cellular ROS accumulation by promoting the degradation of TOC complexes in response to environmental challenge, thereby suppressing the import of photosynthetic proteins. ROS are produced as by-products of photosynthesis, and so SP1 is thought to reduce ROS accumulation by indirectly suppressing photosynthesis under stress conditions.

1.3.3 SP2

The discovery and characterisation of SP1 demonstrated that the UPS directly targets proteins that reside in the chloroplast outer envelope. These proteins are removed from the outer membrane and degraded in the cytosol. However, until recently, little was known as to the mechanisms of protein extraction from the chloroplast.

SP1 was discovered in a screen for extragenic suppressors of the Toc33 null mutant *ppi1*. This screen led to the identification of four mutations at two loci. The second locus was named *SP2*, and its characterisation led to the discovery of a new pathway that mediates the extraction and cytosolic degradation of chloroplast proteins.

SP2 encodes a β-barrel protein that, like SP1, resides in the chloroplast outer envelope membrane (Ling et al., 2019). SP2 is a member of the Omp85 superfamily of membrane proteins (Gentle et al., 2005). All characterised *sp2* mutations can suppress the *ppi1* phenotype with respect to chlorosis, chloroplast development, and TOC protein accumulation. In addition, the *sp2* mutations can supress the *ppi1* phenotype with respect to chloroplast protein import. In this respect, *sp2* mutations are highly similar to *sp1* mutations, which led to the hypothesis that the two encoded proteins might act together in a common pathway. Strong evidence for this hypothesis was provided by a series of careful genetic experiments (Ling et al., 2019). In one of these experiments it was shown that the triple *ppi1 sp1 sp2* mutant displays no enhanced phenotypic suppression (with respect to chlorosis and TOC protein accumulation) relative to the double *ppi1 sp1* and *ppi1 sp2* controls. In addition, it was shown that the singular overexpression of SP1 or SP2 has no obvious phenotypic consequence in the wild-type background; however, the simultaneous overexpression of both genes induces severe chlorosis and developmental stunting. These experiments indicated that SP1 and SP2 function within a common pathway.

SP1 and SP2 physically interact and form a complex that resides in the chloroplast outer envelope membrane. The β-barrel structure of SP2 led to a model in which SP1 ubiquitinates substrates and SP2 functions as a scaffold or channel for their removal from the membrane. In this model, SP2 laterally opens to engulf substrates and provides a channel for their extraction into the cytosol. However, this model does not explain how substrates are actively removed from the chloroplast outer envelope. The membrane extraction of proteins is an energetically costly process, and so the SP1/SP2 complex must associate with at least one other protein that provides the motive force for extraction.

The first clue as to this protein's identity came from a mass spectrometry-based screen for proteins that physically interact with Toc33 (Ling et al., 2019). In this screen, the AAA+ chaperone protein Cdc48 (also known as VCP and p97) was identified. Cdc48

is an extremely well-studied protein that has roles within a very broad range of cellular processes (Baek et al., 2013). Cdc48 is powered by ATP hydrolysis and functions within the disassembly of protein complexes and the extraction of proteins from biological membranes. Importantly, Cdc48 is a key component of the Endoplasmic Reticulum Associated Protein Degradation (ERAD) pathway, a well-characterised pathway that mediates the retrotranslocation, ubiquitination, and degradation of misfolded or redundant ER proteins.

1.3.4 ERAD

The ERAD pathway has been the subject of intense investigation due to its relevance to a number of prevalent and incapacitating human diseases. A detailed overview of ERAD function can be found in the following reviews (Vembar and Brodsky, 2008; Stevenson et al., 2016). In the simplest cases, the pathway proceeds as follows: First, ERAD substrates are recognised by the pathway machinery. This recognition is linked to specific substrate glycosylation patterns. Second, the substrate is guided to a channel protein that facilitates its retrotranslocation into the cytosol. The identity of the channel protein is still unknown. Third, the substrate is transported to the outer edge of the membrane via the channel protein and is ubiquitinated. A small number of E3 ubiquitin ligases have been implicated in ERAD (e.g. Hrd1, gp78) and these can act singularly or cooperatively to polyubiquitinate the substrate. Finally, the substrate is extracted from the membrane and degraded in the cytosol. This step is dependent on Cdc48, which provides the energy for extraction. Substrate ubiquitination is required for Cdc48 recruitment to the membrane.

The ERAD pathway primarily targets misfolded proteins, although in a small number of cases correctly folded proteins are targeted in a regulated fashion. For example, in mammalian cells ERAD has a role in the maintenance of cholesterol homeostasis: HMG-CoA reductase (HMGCR), the rate-limiting enzyme in the cholesterol biosynthesis pathway, is targeted by ERAD under sterol-replete conditions (Menzies et

al., 2018). However, in most cases ERAD mediates the degradation of misfolded proteins. A substrate can be degraded via one of three parallel ERAD pathways, and these are defined by the identity of the E3 ubiquitin ligase involved and by the position of the misfolded lesion relative to the membrane of the endoplasmic reticulum. For example, in yeast, proteins that have misfolded lesions that extend into the cytosol after retrotranslocation are degraded via the action of the Doa10 complex in the ERAD-C pathway (Ruggiano et al., 2014). The other ERAD pathways - ERAD-M and ERAD-L – degrade proteins with lesions present in the ER membrane and ER lumen respectively.

ERAD can be considered the founding member of a small class of similar pathways that function to remove and degrade proteins from membrane-bound compartments. For example, misfolded or redundant proteins are extracted and degraded from the mitochondrial interior via the action of a pathway known as Mitochondrion Associated Degradation (MAD) (Bragoszewski et al., 2017). This pathway is conceptually similar to ERAD and, remarkably, Cdc48 is directly involved. As in ERAD, Cdc48 provides the motive force for substrate extraction.

1.3.5 CHLORAD

In light of various developments within the general cell biology field, the discovery that TOC proteins physically interact with Cdc48 was extremely interesting. Cdc48 was subsequently shown to physically interact with the SP1/SP2 complex, and various lines of genetic and biochemical evidence indicated that it was required for the UPS-dependent removal and degradation of TOC proteins (Ling et al., 2019). Together, the assembly contains an E3 ubiquitin ligase (SP1), a channel protein for substrate retrotranslocation (SP2), and a molecular motor to power substrate retrotranslocation into the cytosol (Cdc48). The pathway was named Chloroplast-Associated Degradation (or CHLORAD) on the basis of its obvious mechanistic and conceptual similarities with ERAD.

1.3.6 How is the CHLORAD pathway regulated?

The discovery of the CHLORAD pathway was a major breakthrough within the chloroplast field as it demonstrated that the UPS extends its reach into the organelle. However, a number of key questions remained unanswered. One of the most pertinent questions related to the regulation of the CHLORAD pathway. The CHLORAD pathway is not constitutively active, and its activity is induced by stress and certain developmental cues (Ling and Jarvis, 2015; Ling et al., 2019). The mechanisms behind this conditional activation are currently unclear.

One clue as to the possible regulation of the CHLORAD pathway came from a large screen for substrates of the small ubiquitin-like modifier protein (SUMO) system in *Arabidopsis* (Elrouby and Coupland, 2010). The SUMO system is a system of post-translational modification that is similar to, and frequently interacts with, the UPS. Intriguingly, Toc159 was found to physically associate with SCE1, the only characterised E2 SUMO conjugating enzyme in *Arabidopsis*. In addition, Toc159 was shown to be SUMOylated in a transgenic *E. coli* strain expressing all the core components of the SUMO system. These findings were interesting in light of the discovery of the CHLORAD pathway, as SUMO has multifaceted roles in cellular responses to stress, it can physically interact with Cdc48 via various Cdc48 cofactors, and crosstalk between SUMO and ubiquitin is common (Kohler et al., 2015; Benlloch and Lois, 2018). As the SUMO system is the major focus of much of this thesis, it will be described in detail next.

1.4 The SUMO system

SUMOylation is an essential form of post-translational modification and it is conserved throughout the eukaryotic domain (Wang and Dasso, 2009; Benlloch and Lois, 2018). SUMO is a ~100 amino acid protein that is structurally similar to ubiquitin. SUMO and ubiquitin are distantly related, and both are conjugated to their substrates via an E1/E2/E3 enzyme cascade. Interestingly, SUMO is just one of a number of ubiquitin-like modifiers found within eukaryotic cells (Schwartz and Hochstrasser, 2003).

The enzymatic architecture of the SUMO system is, as mentioned, broadly similar to that of the UPS. As is the case with the UPS, a small peptide (SUMO) is conjugated to substrates via the action of E1 activating enzymes, E2 conjugating enzymes, and E3 SUMO ligases (Wang and Dasso, 2009). A small class of 'E4' enzymes have a role in SUMO chain synthesis, and a dedicated family of SUMO proteases catalyse the maturation of SUMO peptides, as well as their removal from the substrate (Tomanov et al., 2016).

SUMOylation can modify the activity, localisation, and turnover rate of a protein. The rapid SUMOylation of a large number of proteins is induced by various forms of environmental stress (Benlloch and Lois, 2018). Examples of specific stressors that have been observed to induce SUMO conjugation in plants include heat shock, H_2O_2 exposure, and exposure to ethanol (Kurepa et al., 2003).

1.4.1 The SUMO system in plants

The core components of the *Arabidopsis* SUMO system were identified via genome-wide homology searches and reverse genetic experiments (Kurepa et al., 2003). As is the case in other organisms, the SUMO system is encoded by a very small number of genes compared to the vast number of genes that encode UPS components.

1.4.1.1 SUMO isoforms

The *Arabidopsis* genome contains eight genes that show significant homology with SUMO-encoding genes in other organisms (Kurepa et al., 2003). Of these, only four show detectable expression under standard conditions, and only three have been functionally characterised. All three characterised SUMO proteins are conjugated to exposed lysine residues in target proteins. These lysine residues tend to fall within SUMO 'consensus sequences', which share the same basic amino acid sequence: ΨKXE, where Ψ is a bulky amino acid, K is the targeted lysine residue, X is any amino acid, and E is glutamic acid (Park et al., 2011b).

The most abundant SUMO isoforms are SUMO1 and SUMO2, which share a high degree of amino acid sequence identity (93.5%) and are expressed throughout the plant (Saracco et al., 2007; van den Burg et al., 2010). The genes encoding these two SUMO isoforms are collectively essential; however, the single null mutants *sum1-1* and *sum2-1* do not display any obvious aberrant phenotypes. Although the two genes display substantial functional redundancy, their expression patterns are subtly different. For example, in young leaves *SUMO2* expression is higher in the vasculature, whereas *SUMO1* is more highly expressed in mesophyll cells (van den Burg et al., 2010). However, in mature leaf tissue the expression patterns of the two genes are very similar. SUMO1 and SUMO2 both contain 'classic' SUMO consensus sequences and can themselves be SUMOylated. This SUMOylation leads to the conditional assembly of polySUMO chains (Tomanov et al., 2016).

SUMO3 is more weakly expressed than *SUMO1/2* under standard conditions, and the amino sequence of its encoded protein is significantly divergent (just 48% similar to SUMO1) (van den Burg et al., 2010). Under standard growth conditions *SUMO3* is scarcely expressed; however, its expression is rapidly upregulated under certain circumstances. Null mutants of SUMO3 have been described, and, interestingly, they do not display an obvious aberrant phenotype (van den Burg et al., 2010). However, mutants do display a slight delay in flowering time. SUMO3 has the ability to form polySUMO chains, but it may be a weak chain-builder relative to SUMO1 and SUMO2 (Tomanov et al., 2016).

1.4.1.2 E1 SUMO activating enzymes

E1 'activating' enzymes function within the earliest stage of the SUMO conjugation pathway. Only one complex with E1 function has been described in *Arabidopsis* (Castano-Miquel et al., 2013). The SUMO activating enzyme (SAE) is heterodimeric, comprising a large SAE2 protein and a small SAE1 protein. The large subunit is encoded by a single, essential gene; the null mutant *sae2-1* is non-viable in the homozygous state

(Saracco et al., 2007). The small subunit, SAE1, is encoded by three paralogous genes: *SAE1a*, *SAE1b1* and *SAE1b2*. The three genes actually encode only two distinct products as *SAE1b1* and *SAE1b2* encode proteins of identical amino acid sequence. Both proteins can form a functional complex with SAE2, at least *in vitro*. Whether the different small subunits show any degree of functional specialisation is currently unknown. The three genes show a moderate degree of general functional redundancy. For example, the null mutant *sae1a-1* shows no visible abnormal phenotype under standard growth conditions; however, it does have a reduced ability to form SUMO conjugates under stress conditions (Saracco et al., 2007; Castano-Miquel et al., 2013).

1.4.1.3 E2 SUMO conjugating enzymes

E2 'conjugating' enzymes mediate the transfer of activated SUMO peptides to the substrate. Just one gene - *SCE1* - is known to encode an E2 SUMO ligase in *Arabidopsis* (Saracco et al., 2007). Like other core components of the SUMO system, this gene is essential and null mutations cause embryo lethality in the homozygous state. However, a number of weak alleles that cause strong reductions of *SCE1* expression have been characterised. The *SCE1* locus encodes a small protein (approximately 18kDa) that is abundant throughout the nucleus and cytoplasm (Xiong and Wang, 2013).

1.4.1.4 E3 SUMO ligases

E3 'SUMO ligases' physically interact with the E2 conjugating enzyme and facilitate its interaction with specific substrates. Remarkably, just 2-4 E3 SUMO ligases (depending on definition) have been characterised in *Arabidopsis*. This is in stark contrast to the >1000 E3 ubiquitin ligases thought to be encoded by the *Arabidopsis* genome. Despite the small number of E3 SUMO ligases, a large number of proteins are known to be conditionally SUMOylated.

The two major *Arabidopsis* E3 SUMO ligases are SIZ1 and HPY2 (also known as MMS21) (Miura et al., 2005; Huang et al., 2009). SIZ1 is the better characterised of the two, and it is thought to promote the SUMOylation of a broad range of substrates.

SIZ1 was the first-characterised E3 SUMO ligase and it was identified on the basis of its homology with E3 SUMO ligases found in the baker's yeast and human genomes (Miura et al., 2005). The SIZ1 protein contains a series of conserved domains; these include a SIZ/PIAS-RING domain (SP-RING, known to be essential for E3 SUMO ligase activity in other species and thought to promote physical interactions with E2 SUMO conjugating enzymes), a SUMO interaction motif (SIM; important for SUMO binding), and a nuclear retention sequence. The E3 SUMO ligase activity of SIZ1 has been confirmed *in vitro* and *in vivo*, and a mass spectrometry-based screen indicated that SIZ1 could facilitate the SUMOylation of around 1000 proteins under heat stress conditions (Miura et al., 2005; Rytz et al., 2018).

A series of *siz1* mutants have been identified and characterised. Remarkably, null mutants are viable; however, they are severely stunted and show a very broad range of conditional phenotypes (e.g. sensitivity to freezing (Miura et al., 2007)). A large number of studies have demonstrated the importance of SIZ1 activity in the regulation of various developmental processes. These processes include growth, flowering, phototropism, seed germination, and secondary cell wall formation (Jin et al., 2008; Elrouby, 2015; Kim et al., 2016; Lin et al., 2016; Liu et al., 2019). SIZ1 also has important roles in nutrient assimilation (Park et al., 2011a) and plant responses to various forms of abiotic stress (Miura et al., 2007; Rytz et al., 2018). In addition, SIZ1 is involved in hormone signalling; its activity has been implicated in the regulation of abscisic acid, salicylic acid, and auxin signalling (Lee et al., 2007; Miura et al., 2009; Miura et al., 2011; Zheng et al., 2012).

Relative to SIZ1, HPY2 is poorly understood. Like other E3 SUMO ligases it contains an SP-RING domain; however, unlike SIZ1 it is a simple protein lacking other domains of obvious function. The *HPY2* gene is expressed at a high level throughout the plant and its encoded protein localises to both the nucleus and cytoplasm (Huang et al., 2009; Ishida et al., 2012). The E3 SUMO ligase activity of HPY2 has been confirmed *in vitro* and *in vivo*; however, only a very limited number of HPY2 substrates have been

identified. Surprisingly, a mass spectrometry-based study identified a large number of SIZ1 substrates but failed to identify any HPY2 substrates (Rytz et al., 2018). At time of writing, only two proteins are known to be SUMOylated via the action of HPY2: the cell cycle regulatory protein DPa1, and the chromatin remodelling protein BRAHMA (Liu et al., 2016; Zhang et al., 2017).

Despite the surprisingly small number of known HPY2 substrates, genetic analyses indicate that the gene is of major functional importance. A number of *hpy2* null mutants have been characterised and all display severely stunted growth under standard conditions. The degree of stunting is comparable to that observed in *siz1* mutants, indicating that both E3 SUMO ligases are physiologically important. However, the growth defects, while similar, are not identical. The growth defects observed in *siz1* mutants become more severe with age, with seedlings and young plants showing relatively subtle defects (Ishida et al., 2012). Conversely, *hpy2* mutants show stunting from germination, with particularly severe defects present in root tissues. HPY2 has well-established roles in a number of developmental processes, including cell cycle regulation, stem cell maintenance, and gametophyte development (Huang et al., 2009; Xu et al., 2013; Liu et al., 2014).

1.4.1.5 E4 SUMO ligases

Bioinformatic analyses indicate that the *Arabidopsis* genome contains at least four genes that encode proteins containing the E3 SUMO ligase-associated SP-RING domain (Novatchkova et al., 2012). The first two encode SIZ1 and HPY2, while the third and fourth encode two recently described proteins that have been named PIAL1 and PIAL2 (Tomanov et al., 2016). The *PIAL1* and *PIAL2* genes are both expressed at a moderate level throughout the plant. Null mutants of both genes have been identified and characterised and, surprisingly, neither mutant displays any clear, aberrant phenotype under standard growth conditions. A double *pial1 pial2* null mutant has also been characterised, and, while it shows a subtle growth defect, it does not display

developmental defects of the kind observed in *siz1* and *hpy2* mutants. However, both mutants show altered responses to osmotic and salt stress, implicating the two proteins in plant responses to environmental challenge (Tomanov et al., 2016).

The presence of SP-RING domains in both PIAL1 and PIAL2 implies that both proteins can function as E3 SUMO ligases. However, this was not demonstrated in a series of biochemical experiments (Tomanov et al., 2016). Surprisingly, the double *pial1 pial2* mutant did not show any change in the abundance of SUMO conjugates, either at steady state or in response to heat shock. This finding indicated that PIAL1 and PIAL2 are not typical E3 SUMO ligases. This contention was supported by a second set of experiments in which truncated versions of PIAL1 and PIAL2 were expressed in *E. coli*, purified, and functionally characterised in a series of *in vitro* SUMOylation assays. PIAL1 and PIAL2 were singly added to reaction mixes containing purified SAE, SCE1 and SUMO1, with SUMO1 serving as the SUMO substrate. Remarkably, upon resolving the samples via immunoblotting and probing the membrane with an anti-SUMO1 antibody, the authors observed a very strong increase in the abundance of high-molecular weight smears in the samples incubated with the PIAL proteins. These results implied that both PIAL1 and PIAL2 can promote SUMO chain-building, at least *in vitro*. These findings led the authors to propose that both proteins may function as E4 SUMO ligases (i.e. that they promote SUMO chain-building activity).

1.4.1.6 SUMO proteases

SUMO proteases have two major roles in the SUMO conjugation pathway. SUMO protease function is required for the maturation of nascent SUMO peptides, and also for the removal of SUMO peptides from substrates. A relatively large number of genes encode SUMO proteases in *Arabidopsis*; seven have been functionally characterised so far (Castro et al., 2018). SUMO proteases are not immediately relevant to this study, and so they will only briefly be described here.

All of the characterised *Arabidopsis* SUMO proteases belong to Ub-Like Protease (ULP) family (Castro et al., 2018). Additional families of SUMO proteases have been described in other organisms. The first *Arabidopsis* SUMO protease to be characterised is known as ESD4 (Murtas et al., 2003). ESD4 has a role in the regulation of flowering time, and mutants display reductions in the pool of free SUMO, indicating that ESD4 has a role in the removal of SUMO tags from substrates. ESD4 has two close homologues, ELS1 and ELS2 (Kurepa et al., 2003; Hermkes et al., 2011; Castro et al., 2018). Interestingly, ELS1 seems to have a specialised function. The protein does not appear to mediate the removal of SUMO peptides from substrates, as the *els1-1* null mutant does not show an increase in SUMO conjugation (Hermkes et al., 2011). However, ELS1 does promote the maturation of SUMO peptides *in vitro*. Thus, ELS1 is a specialised SUMO protease with a role in the maturation of SUMO peptides (Hermkes et al., 2011). The second ESD4 homologue, ELS2, is current poorly characterised outside of a small number of bioinformatic analyses (Kurepa et al., 2003; Murtas et al., 2003).

Other SUMO proteases characterised in *Arabidopsis* include: SUMO proteases related to fertility 1 and 2 (SPF1 and SPF2), two proteins that play a poorly-understood role in plant fertility (Liu et al., 2017); and Overly tolerant to salt 1 and 2 (OTS1 and OTS2), two proteins with multifaceted roles in abiotic stress tolerance and hormone signalling (Conti et al., 2008; Bailey et al., 2016).

1.4.2 The SUMO system, Cdc48 and ubiquitin

1.4.2.1 The SUMO system and Cdc48

Cdc48 is a large and sophisticated protein that associates with various cofactors that augment its activity. The two most important Cdc48 cofactors are Ufd1 and Npl4, and Cdc48 is commonly complexed with both proteins. Interestingly, in yeast, Ufd1 can bind to both SUMO and SUMOylated proteins (Nie et al., 2012). This binding has been shown to be physiologically important; for example, the SUMOylation of the DNA double-strand break associated protein Radiation sensitive 52 (Rad52) promotes the recruitment of

Cdc48 to the Rad51/Rad52 recombinase complex, and this activity is important in maintaining genome stability in yeast (Bergink et al., 2013).

1.4.2.2 The SUMO system and the UPS

SUMO and ubiquitin are evolutionary cousins, and they frequently interact to determine the fate of substrate proteins. This crosstalk can be complex, and difficult to predict. SUMO and ubiquitin are both conjugated to exposed lysine residues, and in some cases the SUMOylation of a lysine residue can protect a protein from ubiquitination at the same site. This can promote the stability of a protein by antagonising the action of the UPS. A good example of this behaviour is the case of Drp1, a protein that promotes mitochondrial fission (Harder et al., 2004). Drp1 is recruited from the cytosol to the mitochondrial outer membrane prior to fission. During this recruitment, Drp1 is SUMOylated, and this seems to stabilise the protein by protecting it from the UPS.

SUMOylation can also promote UPS-mediated degradation. A dedicated class of E3 ubiquitin ligases target polySUMO chains: the SUMO-targeted E3 ubiquitin ligases, or STUbLs (for review, see Sriramachandran (2014)). For example, the human STUbL Ring finger protein 4 (RNF4) ubiquitinates mutant CFTR, the protein responsible for cystic fibrosis, and this ubiquitination is triggered by the SUMOylation of the substrate (Ahner et al., 2013).

1.5 Thesis aims and structure

This thesis has two major aims, and both relate directly to the regulation of plastid protein import. The first aim is to explore the relationship between the SUMO system, the CHLORAD pathway, and plastid protein import. In Chapter 2, a series of reverse genetic experiments are described that aim to determine whether there is a genetic link between the SUMO system and plastid protein import. A number of SUMO system mutants are crossed with a pale plastid import mutant, and the resulting double mutants are phenotypically characterised with respect to chlorosis, chloroplast development, and TOC

protein abundance. In Chapter 3, the functional link between the SUMO system and the TOC complex is explored using various molecular approaches. A series of protein-protein interaction studies are described, as are a small number of *in vivo* SUMOylation experiments.

The second aim of this project is to identify new regulators of plastid protein import via forward genetics. In Chapter 4, a forward genetic screen for suppressors of a pale *toc159* (*ppi2*) mutant is described. The new suppressors identified in the screen are phenotypically characterised with respect to chlorosis and general appearance.

2 Genetic experiments reveal a link between plastid protein import and the SUMO system

2.1 Abstract

Chloroplasts import thousands of proteins from the cytosol, and this import is dependent on two cooperative translocons. The outer translocon – the TOC complex – is ubiquitinated and degraded via the action of the CHLORAD pathway. How the CHLORAD pathway is regulated is currently unclear. To determine whether the SUMO system is involved, a series of SUMO mutants were crossed with the pale plastid import mutant *ppi1*. Multiple SUMO mutants could suppress the pale *ppi1* phenotype with respect to various parameters (chlorophyll accumulation, chloroplast development, and TOC protein abundance). In addition, transgenic plants overexpressing SUMO1 or SUMO3 in the *ppi1* background were generated, and these were found to enhance the severity of the *ppi1* phenotype. Taken together, these results demonstrate a strong genetic link between plastid protein import and the SUMO system.

Statement of authorship

Mr Alistair Haslam: Created the following double and triple mutants (*ppi1 sum1-1*; *ppi1 sum2-1*; *ppi1 sum1-1 sum2-1*; *ppi1 sum3-1*; *ppi1 siz1-4*; *ppi1 siz1-5*) and all of the *SUMO1/3* overexpression lines used in this study.

Ms Raman Dhaliwal: Assisted with the transmission electron microscopy (imaging and sample preparation).

2.2 Introduction

Chloroplasts must import thousands of protein species from the cytosol in order to develop and function properly. Such import is dependent on two cooperating complexes that reside in the chloroplast outer and inner membranes: the TOC and TIC complexes. The TOC complex resides in the chloroplast outer membrane and contains three major components: Toc159, Toc33 and Toc75 (Jarvis, 2008; Shi and Theg, 2013). All core components of the TOC complex are conditionally ubiquitinated via the action of SP1, an E3 ubiquitin ligase that resides in the chloroplast outer envelope membrane (Ling et al., 2012). SP1 promotes the turnover of the TOC complex under certain environmental and developmental conditions. This activity has important roles in plant stress tolerance and plastid interconversion events (Ling et al., 2012; Ling and Jarvis, 2015).

SP1 functions within a pathway that meditates the ubiquitination, membrane extraction, and degradation of chloroplast outer membrane proteins (Ling et al., 2019). SP1 physically associates with a second protein, SP2. SP2 is a β-barrel protein that is deeply embedded within the chloroplast outer envelope and is thought to act as a translocon for protein extraction from the membrane. The SP1/SP2 complex physically associates with a third protein, Cdc48. Cdc48 provides the motive force for protein extraction from the membrane. Together, the three proteins – SP1, SP2 and Cdc48 – are the key components of a newly described pathway. This pathway has been named CHLORAD in reference to its similarities with other characterised subcellular pathways.

The regulation of SP1 and the CHLORAD pathway is still poorly understood. The CHLORAD pathway extracts, ubiquitinates and degrades TOC proteins under abiotic stress conditions (Ling and Jarvis, 2015). The pathway's activity must therefore be promoted by various forms of environmental stress; however, how this conditional activation is achieved is currently unclear. The CHLORAD pathway also regulates certain plastid interconversion events (Ling et al., 2012), but, again, how this is achieved is currently unknown.

The activity of the UPS is often modulated by other forms of post-translation modification. One of the best characterised examples of such crosstalk is the interaction between the UPS and the SUMO system. The SUMO system can prevent or promote substrate ubiquitination through various mechanisms, as outlined in detail in Chapter 1. Interestingly, a high-throughput screen for SUMO substrates in *Arabidopsis* identified Toc159 as a potential target of the system (Elrouby and Coupland, 2010). In this study, Toc159 was shown to physically interact with the E2 SUMO conjugating enzyme SCE1, and, in addition, it was SUMOylated by SUMO3 in an *E. coli*-based reconstitution system. SUMOylation is induced by various forms of stress, and so it could be that the CHLORAD-mediated degradation of Toc159 – and potentially other TOC proteins – is promoted by substrate SUMOylation.

The SP1 and SP2 genes were first identified in a forward genetic screen (Ling et al., 2012). Mutations in both loci were shown to act as suppressors of the pale phenotype of Toc33 null mutant *ppi1*. As a first step towards determining whether TOC proteins are SUMOylated, and whether this SUMOylation is physiologically important, a series of reverse genetic experiments were carried out. The pale *ppi1* mutant was crossed with a variety of mutants representing key components of the SUMO pathway, and the resulting double (or, in some cases, triple) mutants were phenotypically characterised. The full catalogue of SUMO mutants used in this study is detailed in Table 2.1. In addition, transgenic lines overexpressing SUMO1 or SUMO3 in the *ppi1* background were generated and phenotypically characterised. In all cases, the mutants or transgenic plants were carefully assessed for phenotypic enhancement or suppression effects, which were taken to indicate a genetic interaction between the SUMO system and plastid protein import.

2.3 Results

2.3.1 The *sce1-4* mutation suppresses the *ppi1* phenotype in a specific manner

The *SCE1* gene encodes the only known E2 SUMO conjugating enzyme in *Arabidopsis*. It therefore follows that if the TOC complex is SUMOylated, SCE1 is highly likely to be involved. To determine whether there is a genetic interaction between SCE1 and plastid protein import, a previously characterised mutant, *sce1-4*, was obtained and crossed with *ppi1*. The *sce1-4* mutant bears a T-DNA insertion slightly upstream of the *SCE1* 5'-UTR and consequently displays a moderate reduction in the expression of *SCE1*, as well as a very mild decrease in the efficiency of heat shock-induced SUMO conjugation (Saracco et al., 2007). Under standard growth conditions, it displays no obvious phenotypic defects. This weak allele was used because null mutations at the *SCE1* locus cause embryo lethality (Saracco et al., 2007).

Intriguingly, *sce1-4* was found to suppress the *ppi1* phenotype with respect to chlorosis (Figure 2.1). The *ppi1 sce1-4* double mutant displayed a moderate increase (+52%) in the abundance of chlorophyll relative to *ppi1* control plants. This effect was highly reproducible and striking given the weak nature of the *sce1-4* allele.

To determine whether the phenotypic suppression observed in *ppi1 sce1-4* was linked to changes in the development of chloroplasts, the chloroplasts of wild type, *ppi1*, and *ppi1 sce1-4* plants were imaged via transmission electron microscopy (TEM) (Figure 2.2). The TEM samples were taken from mature plants that had been grown on soil for 35 days. Remarkably, the chloroplasts of the *ppi1 sce1-4* double mutant were significantly larger (+49%) than the chloroplasts of the *ppi1* control plants. This indicated that the phenotypic suppression observed in *ppi1 sce1-4* was linked to changes in the size and development of chloroplasts.

In *ppi1 sp1* and *ppi1 sp2*, the phenotypic suppression effects are linked to increases in the abundance of TOC proteins (particularly Toc75) (Ling et al., 2012; Ling et al., 2019). To determine whether the phenotypic suppression observed in *ppi1 sce1-4* is linked to similar changes in TOC protein abundance, protein samples were taken from wild-type, *ppi1* and *ppi1 sce1-4* plants. The samples were resolved via immunoblotting and *ppi1 sce1-4* was found to display a clear increase in the abundance of Toc75 relative to the *ppi1* control (Figure 2.3). This indicated that the phenotypic suppression observed in *ppi1 sce1-4* was linked to an increase in the abundance of a key TOC protein.

SCE1 is, as the only known E2 SUMO conjugating enzyme in *Arabidopsis*, a highly promiscuous protein that associates with a very broad range of substrates. It could therefore be argued that the phenotypic suppression effect observed in *ppi1 sce1-4* is potentially linked to processes that are unrelated to chloroplast protein import; that is, that it is a general, pleiotropic effect. To explore this possibility, two control experiments were carried out. First, the *sce1-4* single mutant was phenotypically characterised to determine whether it shows any increase in chlorophyll abundance relative to wild-type plants (Figure 2.4). No effect could be found, which indicated that the phenotypic suppression observed in *ppi1 sce1-4* was a synthetic effect that is specific to the *ppi1* background. Second, *sce1-4* was crossed with the TIC complex-associated mutants *hsp93-V-1* and *tic40-4* (Figure 2.5 and Figure 2.6). Both mutants are chlorotic and have other phenotypic defects that are highly similar to those of *ppi1*. Remarkably, neither the *hsp93-V-1 sce1-4* nor the *tic40-4 sce1-4* double mutant showed any phenotypic suppression relative to single mutant controls. This further supported the conclusion that the suppression effect observed in *ppi1 sce1-4* was a background-specific, synthetic phenotype.

2.3.2 The *siz1-4* mutant suppresses the *ppi1* phenotype in a specific manner

In most cases, SUMOylation is dependent on the action of one or more E3 SUMO ligases. SIZ1 is the best characterised E3 SUMO ligase in *Arabidopsis*. Unfortunately, all

previously described *siz1* mutants display severe dwarf phenotypes and so were not suitable for use in reverse genetic experiments. To overcome this limitation, two new mutants predicted to bear T-DNA insertions within the *SIZ1* locus were identified and characterised. The aim was to identify a weak allele that displays a mild growth defect and could therefore be used in reverse genetic experiments. The two new mutants were named *siz1-4* and *siz1-5*. The *siz1-4* mutant is predicted to bear a T-DNA insertion within the third or fourth exon of *SIZ1*, whereas the *siz1-5* mutant is predicted to bear a T-DNA insertion within the final (or, depending on gene model, penultimate) exon of *SIZ1* (data from the SIGnAL website). Fortuitously, the *siz1-4* mutant was found to display a less severe dwarf phenotype than other characterised *siz1* mutants, and so it was used in most downstream analyses.

The two new *siz1* mutants were crossed with *ppi1* and the resulting double mutants were phenotypically characterised (Figure 2.7). Both *siz1-4* and *siz1-5* could suppress the *ppi1* phenotype with respect to chlorophyll concentration, and this effect was apparent from early stages of development (at which point the *siz1* mutants did not show severe growth defects and so could be better compared with control plants). The *ppi1 siz1-4* double mutant showed a remarkably strong increase in chlorophyll concentration; at maturity, the double mutant displayed an almost wild-type level of chlorophyll accumulation.

To determine whether the phenotypic suppression effect observed in *ppi1 siz1-4* was linked to changes in TOC protein abundance, protein samples were taken from the double mutant alongside wild-type and *ppi1* controls (Figure 2.8). Interestingly, the *ppi1 siz1-4* double mutant showed a significant increase in the abundance of both Toc75 and Toc159, relative to *ppi1*.

In an effort to demonstrate that the suppression effect observed in *ppi1 siz1-4* was a specific, synthetic phenotype, the *siz1-4* single mutant was phenotypically characterised with respect to chlorophyll concentration (Figure 2.9). Measurements were taken after

approximately 5 weeks of growth (as in the *ppi1 siz1-4* analysis). Remarkably, the mutant did not show any change in chlorophyll abundance relative to a wild type control. This is notable given that the *ppi1 siz1-4* double mutant displayed a very large increase in chlorophyll concentration relative to a *ppi1* single mutant control (>+100%).

2.3.3 The *sum1-1* and *sum2-1* mutants additively suppress the *ppi1* phenotype

There are three characterised SUMO proteins in *Arabidopsis*: SUMO1, SUMO2 and SUMO3 (see Chapter 2). SUMO1 and SUMO2 are highly similar in terms of amino acid sequence (93.5% sequence identity) and expression pattern (Saracco et al., 2007; van den Burg et al., 2010). Null mutants of both SUMO1 and SUMO2 have been described, and neither shows any clear aberrant phenotype under standard growth conditions (Saracco et al., 2007). However, the double null mutant *sum1-1 sum2-1* is non-viable in the homozygous state. The *sum1-1* and *sum2-1* mutants were both crossed with *ppi1* and the resulting plants were phenotypically characterised. In addition, to circumvent the problem of redundancy, a triple *ppi1 sum1-1 sum2-1* mutant was sought. As the homozygous double *sum1-1 sum2-1* mutant is non-viable, a line homozygous with respect to *ppi1* and *sum2-1*, but heterozygous with respect to *sum1-1*, was identified from a large pool of segregating plants. The genotype of each triple mutant was confirmed via a PCR-based method. The triple mutant was grown alongside the double mutants and relevant control plants, and the phenotypes were recorded (Figure 2.10). While both *sum1-1* and *sum2-1* individually could only weakly suppress the *ppi1* phenotype with respect to chlorophyll accumulation (effect sizes were +14% and +17%, respectively), the triple mutant showed a robust suppression effect, displaying a moderate increase in chlorophyll accumulation (+>40%) relative to a *ppi1* control. This was striking given that the triple mutant still retained strong residual SUMOylation (van den Burg et al., 2010), and implied that the two major SUMO proteins are involved in TOC SUMOylation.

2.3.4 SUMO1 overexpression enhances the *ppi1* phenotype

To complement the reverse genetic experiments, transgenic plants overexpressing SUMO1 in the *ppi1* background were generated. The *SUMO1* coding sequence was cloned into pB2GW7, a vector carrying a strong, constitutive promotor (cauliflower mosaic virus 35S) upstream of the cloning site (Karimi et al., 2002). The resulting construct was introduced into the *ppi1* background via floral dipping. Two lines carrying a single insert in the homozygous state were identified and characterised (Figure 2.11). The overexpression of *SUMO1* was confirmed in both lines via semi-quantitative RT-PCR (Figure 2.11). Both lines displayed an accentuation of the *ppi1* phenotype: the plants were significantly smaller and showed more severe chlorosis (-10% and -13%) relative to a *ppi1* control. This finding complemented the results of the *sum1-1* and *sum2-1* analysis (Section 2.3.3) and suggested that SUMO might act to negatively regulate plastid protein import, perhaps through the modulation of the CHLORAD pathway.

2.3.5 The *sum3-1* mutant cannot suppress the *ppi1* phenotype

SUMO3 is dissimilar to SUMO1 and SUMO2 in terms of amino acid sequence (43% amino acid sequence identity shared with SUMO1) and expression pattern (Kurepa et al., 2003; van den Burg et al., 2010). It was also the SUMO isoform found to physically interact with Toc159 in a high-throughput screen for SUMO substrates in *Arabidopsis* (Elrouby and Coupland 2010). Interestingly, SUMO3 is not essential for plant viability: null mutants do not show any major growth defects under standard conditions (van den Burg et al., 2010). A previously characterised null mutant, *sum3-1*, was obtained for use in genetic experiments (van den Burg et al., 2010). This mutant carries a T-DNA insertion in the third exon of *SUMO3* which completely abrogates transcription. The *sum3-1* mutant was crossed with *ppi1* and the resulting double mutant was characterised (Figure 2.12). Interestingly, the double mutant displayed no obvious phenotypic change relative to a single *ppi1* control.

2.3.6 SUMO3 overexpression enhances the *ppi1* phenotype

To complement the preceding genetic experiment, transgenic lines overexpressing SUMO3 in the *ppi1* background were generated and phenotypically studied. The *SUMO3* cDNA was cloned into the pB2GW7 vector and the resulting construct was introduced into the *ppi1* background (Karimi et al., 2002). A line bearing a single, homozygous insert was identified and characterised (Figure 2.13). The overexpression of *SUMO3* was confirmed in the line via semi-quantitative RT-PCR (Figure 2.13). The plants displayed significant phenotypic enhancement with respect to both size and chlorophyll accumulation (~-40%) relative to a *ppi1* control. This finding is consistent with a role for SUMO3 in the regulation of plastid import and, possibly, the CHLORAD pathway. However, the fact that the *sum3-1* mutant could not suppress *ppi1* (Section 2.4.5) is surprising.

2.4 Discussion

The experiments described in this chapter demonstrated a strong genetic link between the SUMO system and plastid protein import. Mutants representing various components of the SUMO system (the E2 SUMO conjugating enzyme, an E3 SUMO ligase, and two SUMO isoforms) were found to suppress the pale *ppi1* phenotype. The inverse effect was observed when SUMO1 and SUMO3 were overexpressed in the *ppi1* background. In addition, the phenotypic suppression observed in the double mutants was linked to increases in chloroplast size and TOC protein abundance.

The genetic experiments were inspired by the forward genetic screen in which SP1 and SP2 were discovered (Ling et al., 2012; Ling et al., 2019). In this screen, *sp1* and *sp2* mutants were found to suppress *ppi1* and thus revealed a genetic link between the UPS and plastid protein import. It was therefore reasoned that if SUMO regulates plastid import and/or the CHLORAD pathway, SUMO system mutants would also modify the phenotype of *ppi1*. However, the 'direction' of any such modification – i.e. phenotypic enhancement vs. suppression – was completely unknown. During the early stages of the project it was hypothesised that SUMO might antagonise the action of SP1 by shielding the exposed lysine residues of TOC proteins from ubiquitination. This model was based on the case of Drp1, a mitochondrial fission factor that is stabilised in the outer mitochondrial membrane by SUMOylation (Harder et al., 2004). If this model were correct, SUMO mutants would be expected to enhance the severity of the *ppi1* phenotype.

However, as the first genetic experiments were completed, it became obvious that the double mutants displayed phenotypic suppression, not enhancement. This ruled out the first model of SUMO function at the TOC complex. Intriguingly, almost all of the double mutants phenotypically resembled *ppi1 sp1* and *ppi1 sp2* with respect to chlorophyll accumulation, chloroplast size, and TOC protein abundance. Similarly, the *ppi1* SUMO1 OX and *ppi1* SUMO3 OX transgenic plants phenotypically resembled *ppi1* SP1 OX and *ppi1* SP2 OX plants, in that they display an enhanced *ppi1* phenotype (the

plants were smaller and paler). These findings were consistent with an alternative model in which SUMO acts to promote the action of the CHLORAD pathway, perhaps in response to stress. This idea is extensively explored in Chapter 5. To robustly support the link between the SUMO and CHLORAD systems, two triple mutants were sought: *ppi1 sp1 sce1-4* and *ppi1 sp2 sce1-4*. The triple mutants were expected to display no increase in phenotypic suppression relative to *ppi1 sp1* and *ppi1 sp2* controls, which would indicate that SUMO and the CHLORAD machinery function within a common pathway. Unfortunately, the identification and analysis of these lines was not finalised by the end of this project.

The experiments described in this chapter not only demonstrated a genetic link between the SUMO system and plastid protein import, but they also hinted at the identity of the specific SUMO system components that are involved. The E3 SUMO ligase SIZ1 was implicated, but, unfortunately, the other major E3 SUMO ligase HPY2 was not assessed in this study. This was due to an absence of suitable mutants (i.e., hypomorphic alleles that do not cause severe growth defects) and the relative scarcity of confirmed HPY2 substrates (see Chapter 1). In hindsight, this omission was regrettable, given that HPY2 shows cytosolic localisation (Huang et al., 2009) and features in the next chapter. The two E4 SUMO ligases PIAL1 and PIAL2 (Tomanov et al., 2016) were also not assessed in this study, which is regrettable as polySUMO chains are targeted by some classes of E3 ubiquitin ligases (Sriramachandran and Dohmen, 2014).

The three principal SUMO isoforms were assessed in this study. There was a slight incongruence between the reverse genetic and transgenic experiments. Analyses of double and triple mutants implied that SUMO1 and SUMO2, but not SUMO3, regulate plastid protein import. However, the phenotypic enhancement effect observed in the *ppi1* SUMO3 OX transgenic plants was much stronger than that observed in the *ppi1* SUMO1 OX plants. In addition, the *ppi1* SUMO3 OX plants were severely stunted, and this stunting may be a synthetic effect – transgenic plants overexpressing SUMO3 in the wild-

type background did not show any growth defects in one study (van den Burg et al., 2010). The absence of a suppression effect in *ppi1 sum3-1* is puzzling. One possibility is that SUMO1 and SUMO2 may compensate for SUMO3. At least two studies have found that there is considerable overlap between SUMO1 and SUMO3 substrates (Budhiraja et al., 2009; Elrouby and Coupland, 2010). In addition, *SUMO1* and *SUMO2* are much more strongly expressed than *SUMO3* (van den Burg et al., 2010), so even if the two proteins show a lower affinity for TOC substrates, this could be compensated for by the large pool of the two proteins.

The genetic experiments described in this chapter have a potential weakness: all of the SUMO mutants that were studied are expected to have complex molecular phenotypes due to the extreme promiscuity of the key SUMO system proteins (Elrouby and Coupland, 2010). Linking the suppression effects to a specific process (i.e., plastid protein import) is therefore challenging. However, two sets of experiments were undertaken to address this problem. In the first, the single *sce1-4* and *siz1-4* mutants were phenotypically assessed and found to display no increase in chlorophyll accumulation. This implied that the suppression effects observed in the double mutants were synthetic phenotypes specific to the *ppi1* background. The *siz1-4* result was particularly convincing, as *ppi1 siz1-4* showed a dramatic increase in chlorophyll accumulation, almost completely suppressing the *ppi1* phenotype with respect to chlorosis. In the second set of experiments, *sce1-4* was crossed with two additional chlorotic mutants: *tic40-4* and *hsp93-V-1*. The resulting double mutants did not display any increase in chlorophyll accumulation. These were striking results, as the two mutants are associated with the TIC complex and are highly similar to *ppi1* in terms of function and appearance (Kovacheva et al., 2005). These experiments strongly argued that the observed suppression effects are synthetic and highly specific to the *ppi1* background.

In future work, the phenotypic suppression and enhancement effects observed in the various double mutants and transgenic plants should be directly linked to changes in

plastid import efficiency. The *ppi1 sp1* and *ppi1 sp2* mutants both show an increase in plastid protein import efficiency, which fully explains the phenotypic suppression effect in the two lines (Ling et al., 2012; Ling et al., 2019). Given that most of the *ppi1*/SUMO system double mutants resemble *ppi1 sp1* and *ppi1 sp2* in terms of chlorophyll accumulation, chloroplast development, and TOC protein abundance, it is highly likely that a similar effect on plastid import would be observed. The genetic experiments outlined in this chapter heavily imply a direct link between the SUMO system and plastid protein import but, due to time constraints and technical challenges, this has not been directly demonstrated.

The results described in this chapter demonstrate a strong and specific genetic link between the SUMO system and plastid protein import, and that the E3 SUMO ligase SIZ1 and all three SUMO isoforms may be involved. The results also imply that SUMOylation regulates chloroplast development and the stability of TOC proteins. Taken together, these findings are consistent with a model in which SUMOylation promotes the turnover of TOC proteins, perhaps in concert with the CHLORAD pathway.

2.5 Figures and tables

Table 2.1

A catalogue of the SUMO mutants used in this study.

Allele name	AGI number	Gene name	Catalogue ID	Gene function	Allele type (null, weak)	Reference
sum1-1	AT4G26840	SUMO1	SAIL_296_C12	SUMO encoding (major isoform)	Null	(Saracco et al., 2007)
sum2-1	AT5G55160	SUMO2	SALK_129775	SUMO encoding (major isoform)	Null	(Saracco et al., 2007)
sum3-1	AT5G55170	SUMO3	SM_3.2707	SUMO encoding (minor isoform)	Null	(van den Burg et al., 2010)
sce1-4	AT3G57870	SCE1	SALK_006164	E2 SUMO conjugating enzyme	Weak	(Saracco et al., 2007)
siz1-4	AT5G60410	SIZ1	SAIL_805_A10	E3 SUMO ligase	Presumed weak	-
siz1-5	AT5G60410	SIZ1	SALK_111280	E3 SUMO ligase	Presumed null	-

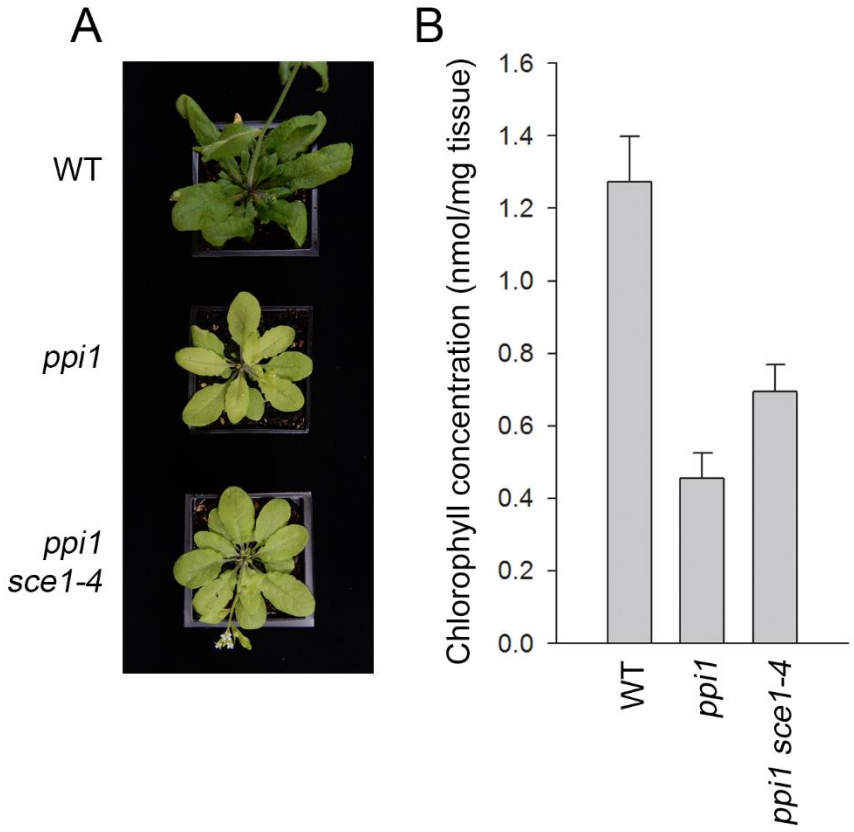


Figure 2.1

Suppression of the *ppi1* phenotype by the E2 SUMO conjugating enzyme mutation *sce1-4.*

(A) Visual appearance of the *ppi1 sce1-4* double mutant and appropriate control plants. Representative plants are shown. The plants were grown on soil together under identical conditions and were approximately 5 weeks old at the time of photography.

(B) Quantification of chlorophyll concentration in the *ppi1 sce1-4* double mutant. Measurements were taken from rosette tissue on the day of photography (A). A SPAD meter was used. Error bars indicate standard deviation from the mean (n = 7). There were significant differences between the *ppi1* and *ppi1 sce1-4* plants (Two-tailed t-test, unpaired samples, T = 6.15, p = 0.000049).

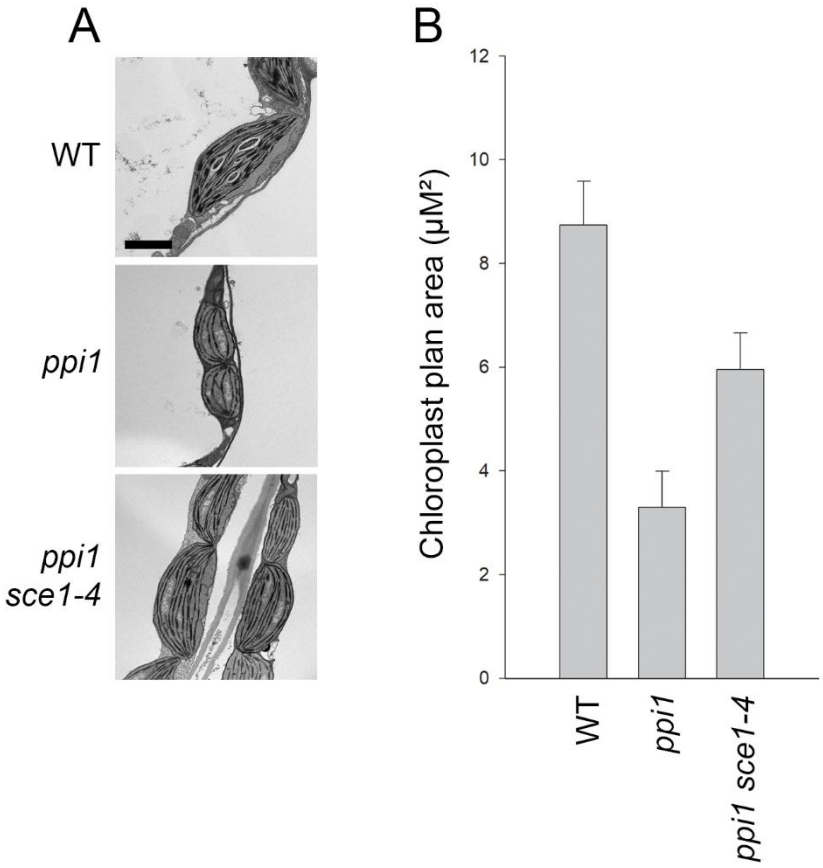


Figure 2.2

The *ppi1 sce1-4* double mutant displays an increase in chloroplast size.

(A) Transmission electron microscopy images of chloroplasts in *ppi1 sce1-4* and appropriate control plants. Samples for analysis were taken from rosette leaves of plants grown on soil for approximately 4 weeks. Representative images are shown. Scale bar = 2 µM.

(B) Quantification of chloroplast size in *ppi1 sce1-4* and appropriate control plants. The image set from which the images in panel A were derived was analysed. Three plants were imaged per genotype, and approximately 10 micrographs were taken from each plant. Error bars indicate standard deviation from the mean (n = 3). There were significant differences between the *ppi1* and *ppi1 sce1-4* plants (Two-tailed t-test, unpaired samples, T = 4.65, p = 0.009674).

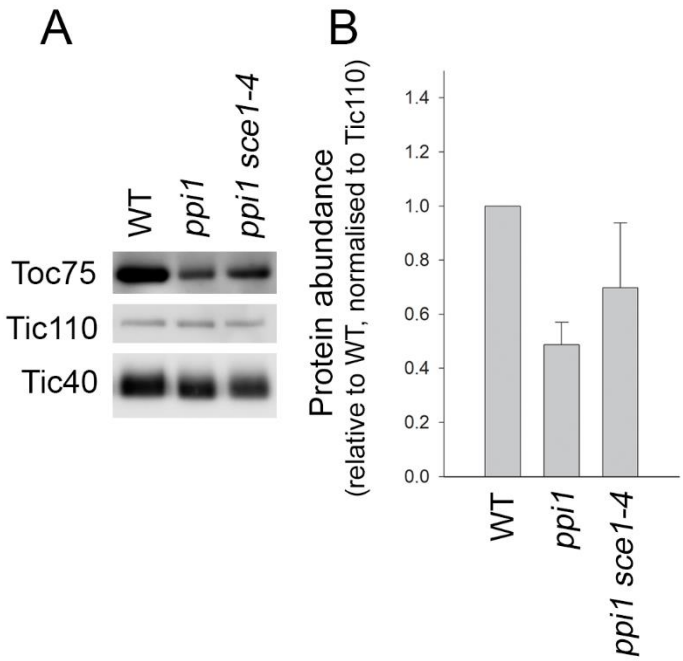


Figure 2.3

The *ppi1 sce1-4* double mutant displays an increase in the abundance of Toc75.

(A) Immunoblot analysis of Toc75 abundance in the indicated genotypes. Protein samples for analysis were taken from mature rosette leaves of plants grown on soil for approximately 4 weeks. Loading was normalised against the intensity of Tic110, a chloroplast control protein. Tic40 was included as an additional loading control.

(B) Quantification of Toc75 abundance in the indicated genotypes. Immunoblot band intensities were quantified using ImageJ. The values shown represent Toc75:Tic110 band intensity ratios, normalised against the wild-type control value. Error bars represent standard error of the mean (n = 2).

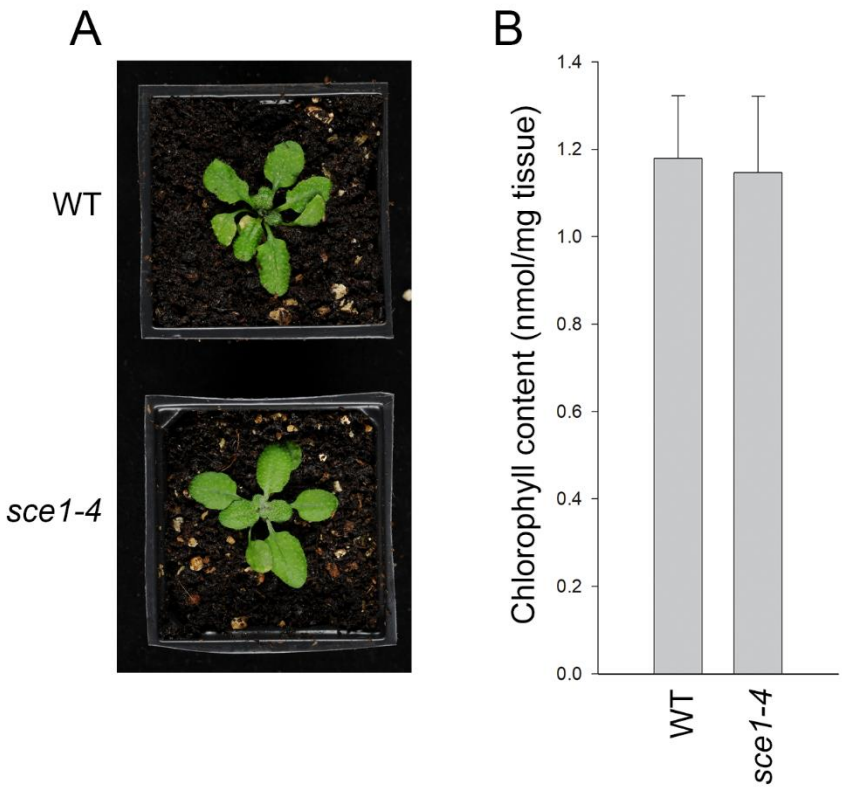


Figure 2.4

The *sce1-4* single mutant does not display an increase in chlorophyll accumulation.

(A) Visual appearance of the *sce1-4* single mutant and wild type control. Representative plants are shown. The plants were grown on soil together under identical conditions and were approximately 4 weeks old at the time of photography.
(B) Quantification of chlorophyll concentration in the *sce1-4* single mutant. Measurements were taken from rosette tissue on the day of photography (A). A SPAD meter was used. Error bars indicate standard deviation from the mean (n = 7). There were no significant differences between the *sce1-4* and wild type plants (Two-tailed t-test, unpaired samples, T = 0.38, p = 0.708484).

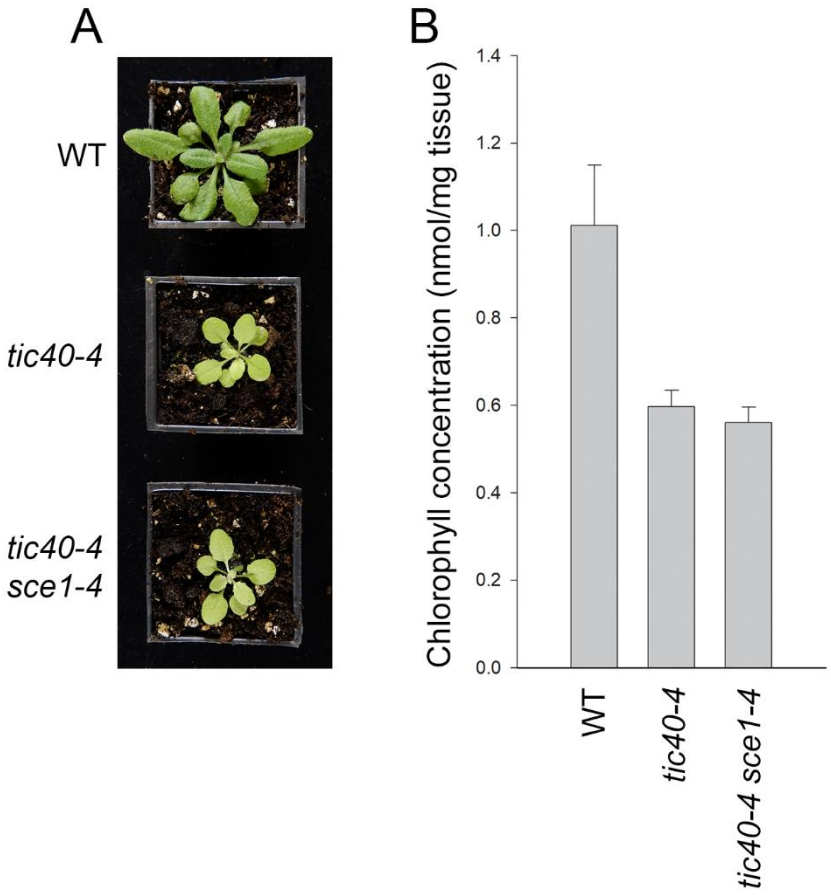


Figure 2.5

The *sce1-4* mutant does not suppress the pale phenotype of plastid import mutant *tic40-4*.

(A) Visual appearance of the *tic40-4 sce1-4* double mutant and appropriate control plants. Representative plants are shown. The plants were grown on soil together under identical conditions and were approximately 4 weeks old at the time of photography.

(B) Quantification of chlorophyll concentration in the *tic40-4 sce1-4* double mutant. Measurements were taken from rosette tissue on the day of photography (A). A SPAD meter was used. Error bars indicate standard deviation from the mean (n = 5). There were no significant differences between the *tic40-4* and *tic40-4 sce1-4* plants (Two-tailed t-test, unpaired samples, T = 1.25, p = 0.280106).

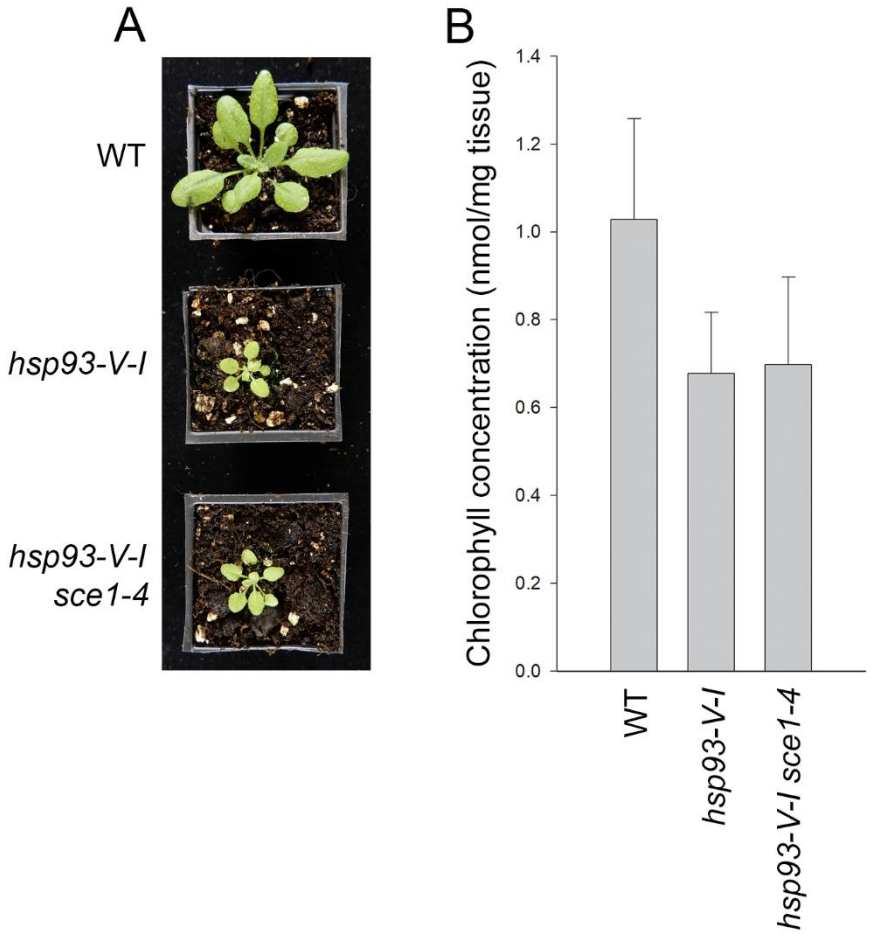


Figure 2.6

The *sce1-4* mutant does not suppress the pale phenotype of plastid chaperone mutant *hsp93-V-I*.

(A) Visual appearance of the *hsp93-V-I sce1-4* double mutant and appropriate control plants. Representative plants are shown. The plants were grown on soil together under identical conditions and were approximately 4 weeks old at the time of photography.
(B) Quantification of chlorophyll concentration in the *hsp93-V-I sce1-4* double mutant. Measurements were taken from rosette tissue on the day of photography (A). A SPAD meter was used. Error bars indicate standard deviation from the mean (n = 4). There were no significant differences between the *hsp93-V-I* and *hsp93-V-I sce1-4* plants (Two-tailed t-test, unpaired samples, T = 0.18, p = 0.860702).

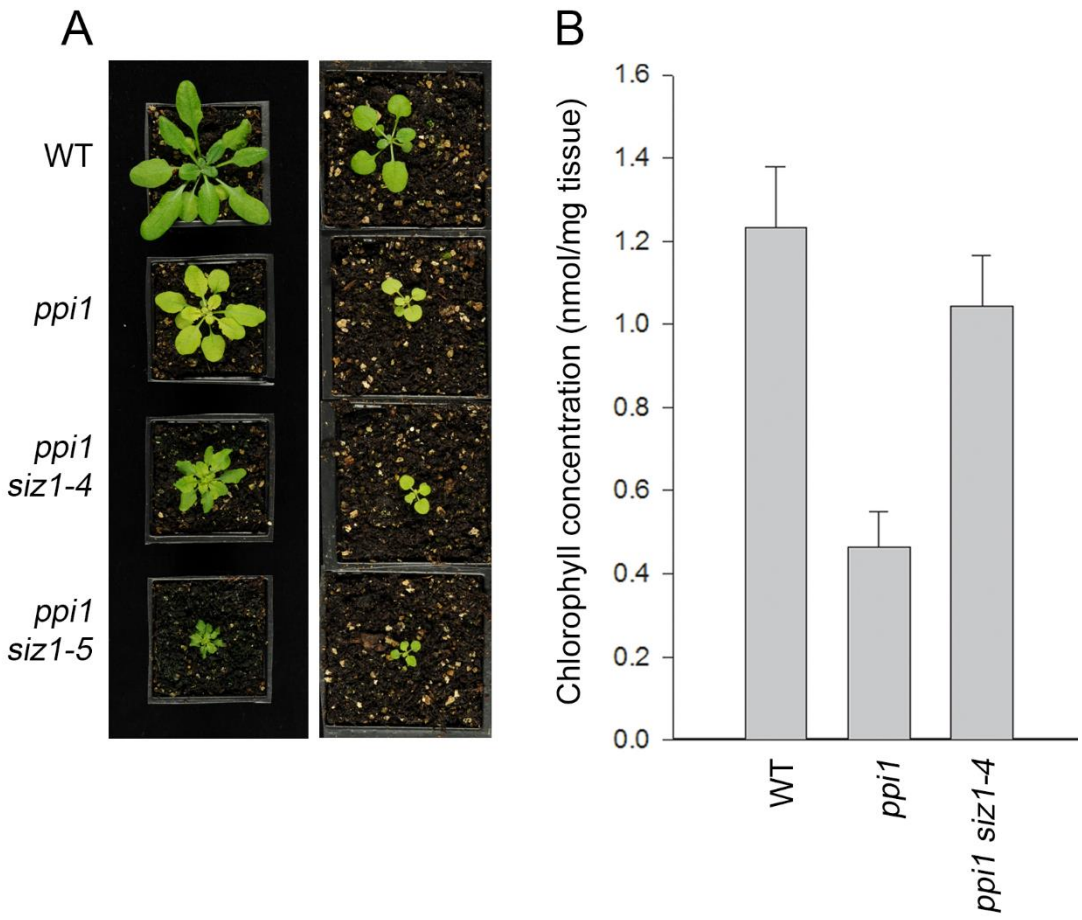


Figure 2.7

Suppression of the *ppi1* phenotype by the E3 SUMO ligase mutations *siz1-4* and *siz1-5*.

(A) Visual appearance of the *ppi1 siz1-4* and *ppi1 siz1-5* double mutants and appropriate control plants. Representative plants are shown. The plants were grown on soil together under identical conditions and were approximately 5 weeks old (left hand panel) or 3 weeks old (right hand panel) at the time of photography.
(B) Quantification of chlorophyll concentration in the *ppi1 siz1-4* double mutant. Measurements were taken from rosette tissue on the day of photography (A). A SPAD meter was used. Error bars indicate standard deviation from the mean (n = 8). There were significant differences between the *ppi1* and *ppi1 siz1-4* plants (Two-tailed t-test, unpaired samples, T = 11.01, $p < 0.00001$).

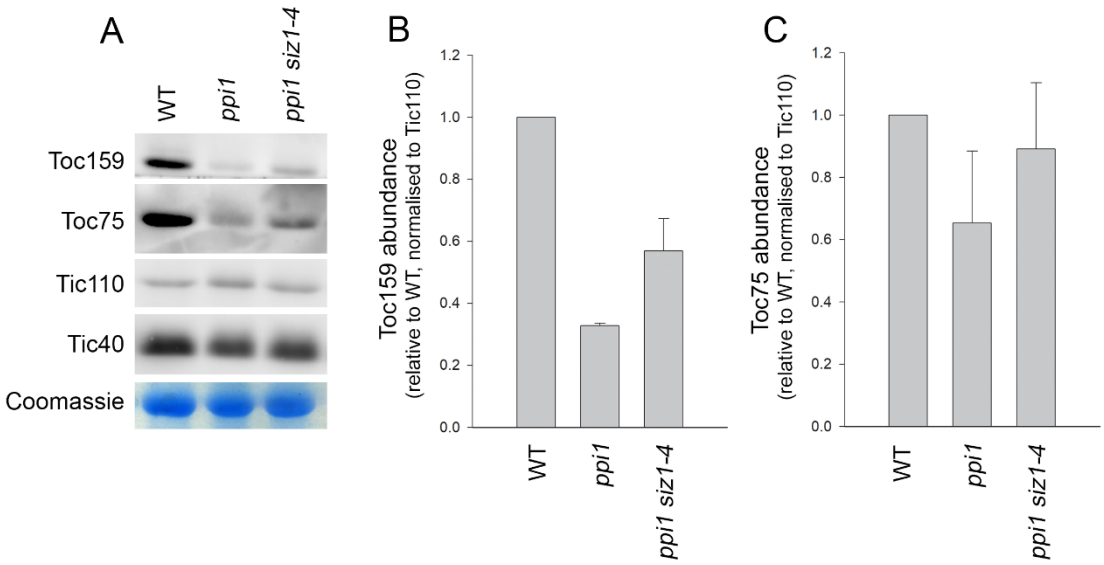


Figure 2.8

The *ppi1 siz1-4* double mutant displays an increase in the abundance of Toc75 and Toc159.

(A) Immunoblot analysis of Toc75 and Toc159 abundance in the indicated genotypes. Protein samples for analysis were taken from mature rosette leaves of plants grown on soil for approximately 4 weeks. Loading was normalised against the intensity of Tic110, a chloroplast control protein. Tic40 and a Coomassie stain served as additional loading controls.
(B) Quantification of Toc159 abundance in the indicated genotypes. Immunoblot band intensities were quantified using ImageJ. The values shown represent Toc159:Tic110 band intensity ratios, normalised against the wild-type control value. Error bars represent standard error of the mean (n = 2).
(C) Quantification of Toc75 abundance in the indicated genotypes. Immunoblot band intensities were quantified using ImageJ. The values shown represent Toc75:Tic110 band intensity ratios, normalised against the wild-type control value. Error bars represent standard error of the mean (n = 2).

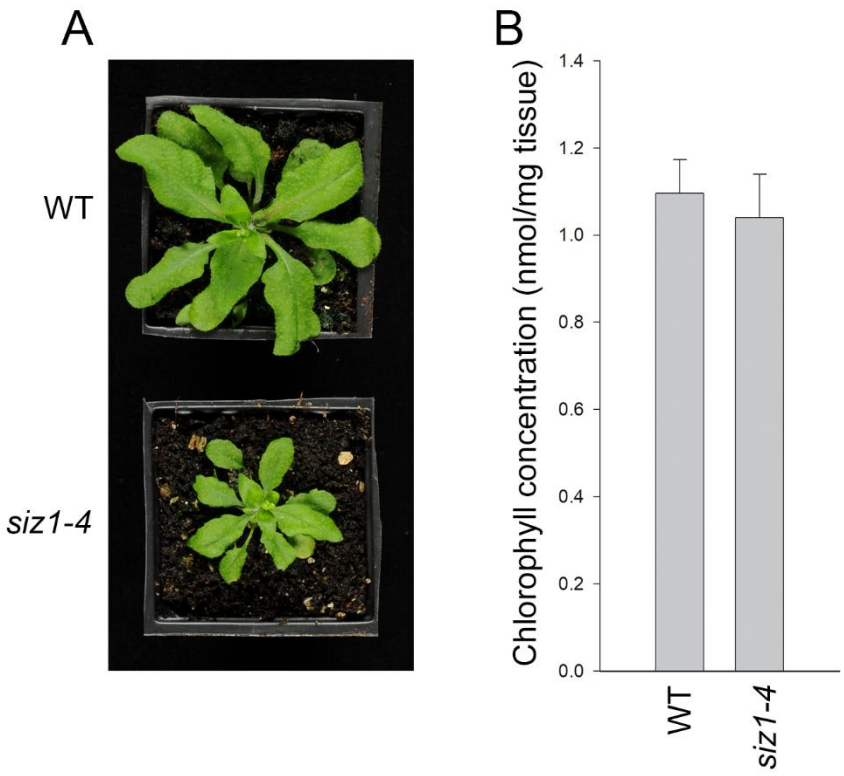


Figure 2.9

The *siz1-4* single mutant does not display an increase in chlorophyll accumulation.

(A) Visual appearance of the *siz1-4* single mutant and wild type control. Representative plants are shown. The plants were grown on soil together under identical conditions and were approximately 5 weeks old at the time of photography.
(B) Quantification of chlorophyll concentration in the *siz1-4* single mutant. Measurements were taken from rosette tissue on the day of photography (A). A SPAD meter was used. Error bars indicate standard deviation from the mean (n = 4). There were no significant differences between the *siz1-4* and wild type plants ($T = 0.96$, $p = 0.370055$).

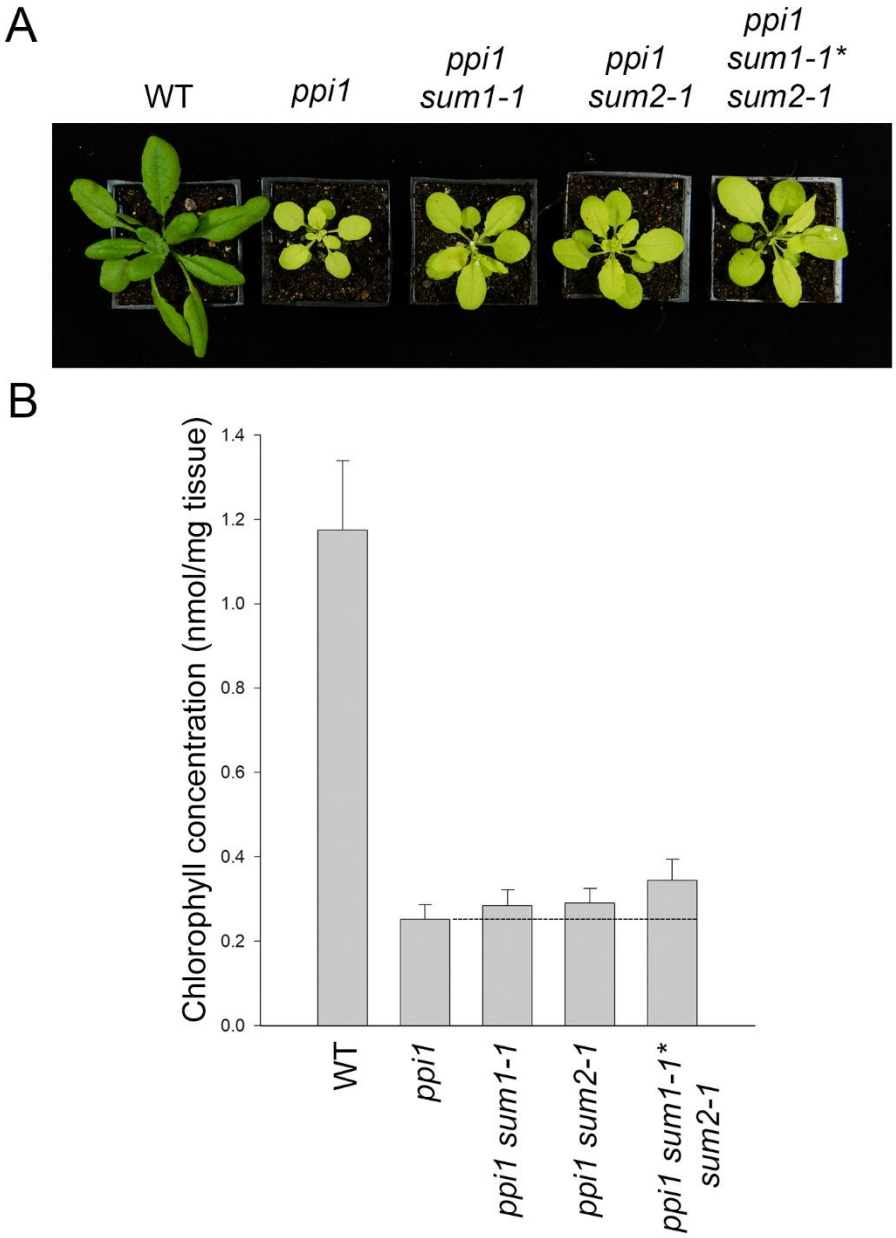


Figure 2.10

Additive suppression of the *ppi1* phenotype by the SUMO isoform mutations *sum1-1* and *sum2-1*

(A) Visual appearance of the double and triple mutants and appropriate control plants. Representative plants are shown. The plants were grown on soil together under identical conditions and were approximately 4 weeks old at the time of photography. The asterisk indicates that the triple mutant is heterozygous with respect to the *sum1-1* mutation.

(B) Quantification of chlorophyll concentration in the double and triple mutants. Measurements were taken from rosette tissue on the day of photography (A). A SPAD meter was used. Error bars represent standard deviation from the mean (*n* = ≥20). The dotted line represents the *ppi1* value. T-tests were performed (unpaired, two-tailed). There was a significant difference between *ppi1* and *ppi1*

sum1-1 (T = 3.979; *p* = 0.0002). There was also a significant difference between *ppi1* a*nd ppi1 sum2-1* (T = 5.490; *p* = <0.00001). There was also a significant difference between *ppi1* and *ppi1 sum1-1* sum2-1* (T = 8.160; *p* = <0.00001).

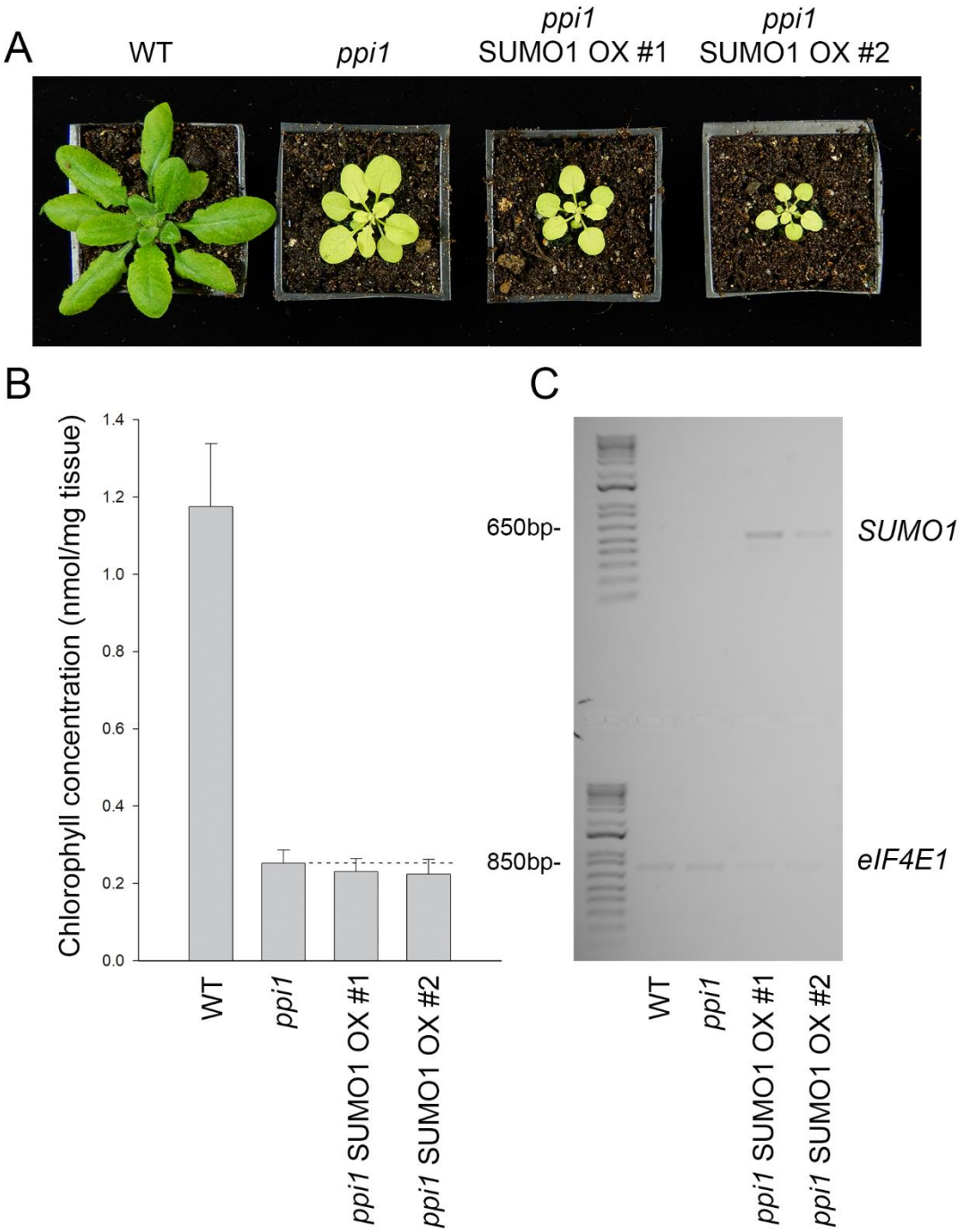


Figure 2.11

Accentuation of the *ppi1* phenotype by SUMO1 overexpression.

(A) Visual appearance of the *ppi1* SUMO1 OX transformants. Representative plants are shown. The plants were grown on soil together under identical conditions and were approximately 5 weeks old at the time of photography. Two independent, single-insert SUMO1 OX lines are shown.

(B) Quantification of chlorophyll concentration in the *ppi1* SUMO1 OX transformants. Measurements were taken from rosette tissue on the day of photography (A). Error bars indicate standard deviation from the mean ($n \geq 13$). The dotted line represents the *ppi1* value. T-tests were performed (unpaired, two-tailed). There

was a significant difference between *ppi1* and *ppi1* SUMO1 OX 3-6 ($T = 2.40$; $p < 0.05$). There was also a significant difference between *ppi1* and *ppi1* SUMO1-OX 5-6 ($T = 2.49$; $p < 0.05$).

(C) Analysis of 35S:SUMO1 transgene expression in the two SUMO1 OX transformants by semi-quantitative RT-PCR. Total RNA samples isolated from rosette tissue of plants that were approximately 4 weeks old were used as a template for cDNA synthesis. Amplifications employed primers specific to *SUMO1* and to the reference gene *eIF4E1*. A limited number of amplification cycles (25) was employed in each case to prevent saturation, and the products were analysed via agarose gel electrophoresis. Amplicon sizes are indicated by the DNA ladder (see Chapter 6).

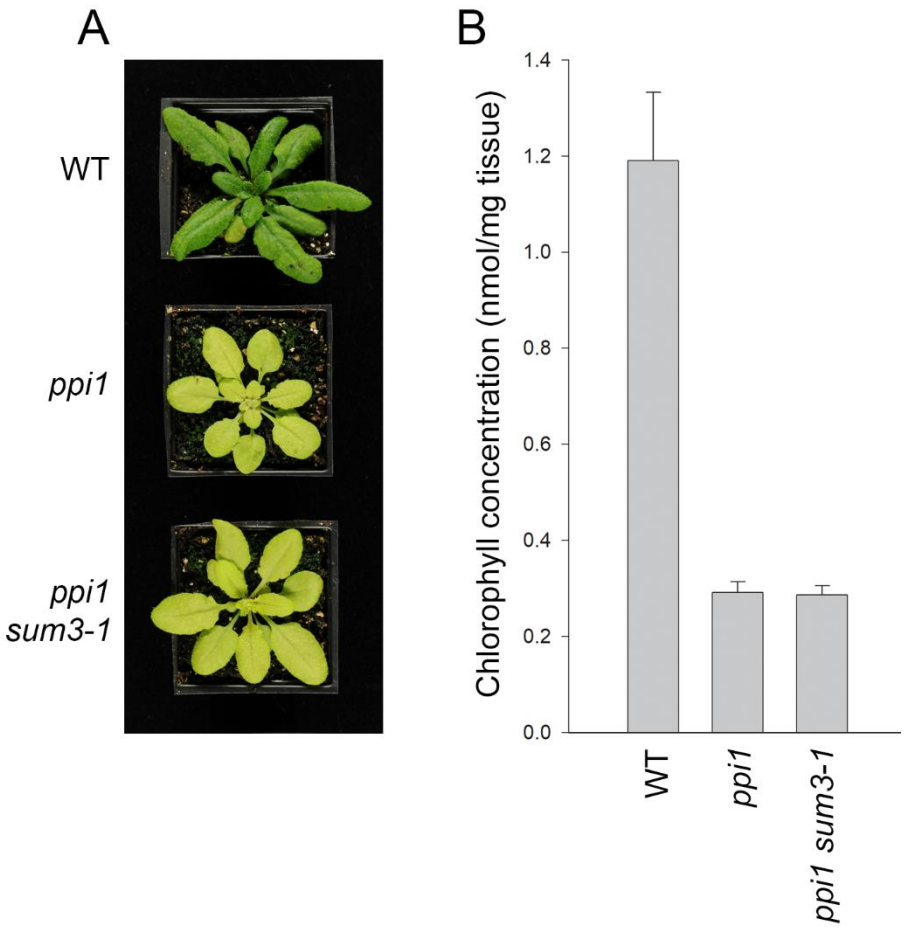


Figure 2.12

The *sum3-1* mutation does not suppress the pale *ppi1* phenotype.

(A) Visual appearance of the *ppi1 sum3-1* double mutant. Representative plants are shown. The plants were grown on soil together under identical conditions and were approximately 4 weeks old at the time of photography.
(B) Quantification of chlorophyll concentration in the *ppi1 sum3-1* double mutant. Measurements were taken from rosette tissue on the day of photography (A). A SPAD meter was used. Error bars indicate standard deviation from the mean (n = 10). There were no significant differences between the *ppi1* and *ppi1 sum3-1* plants (Two-tailed t-test, unpaired samples, T = 0.54, p = 0.59407).

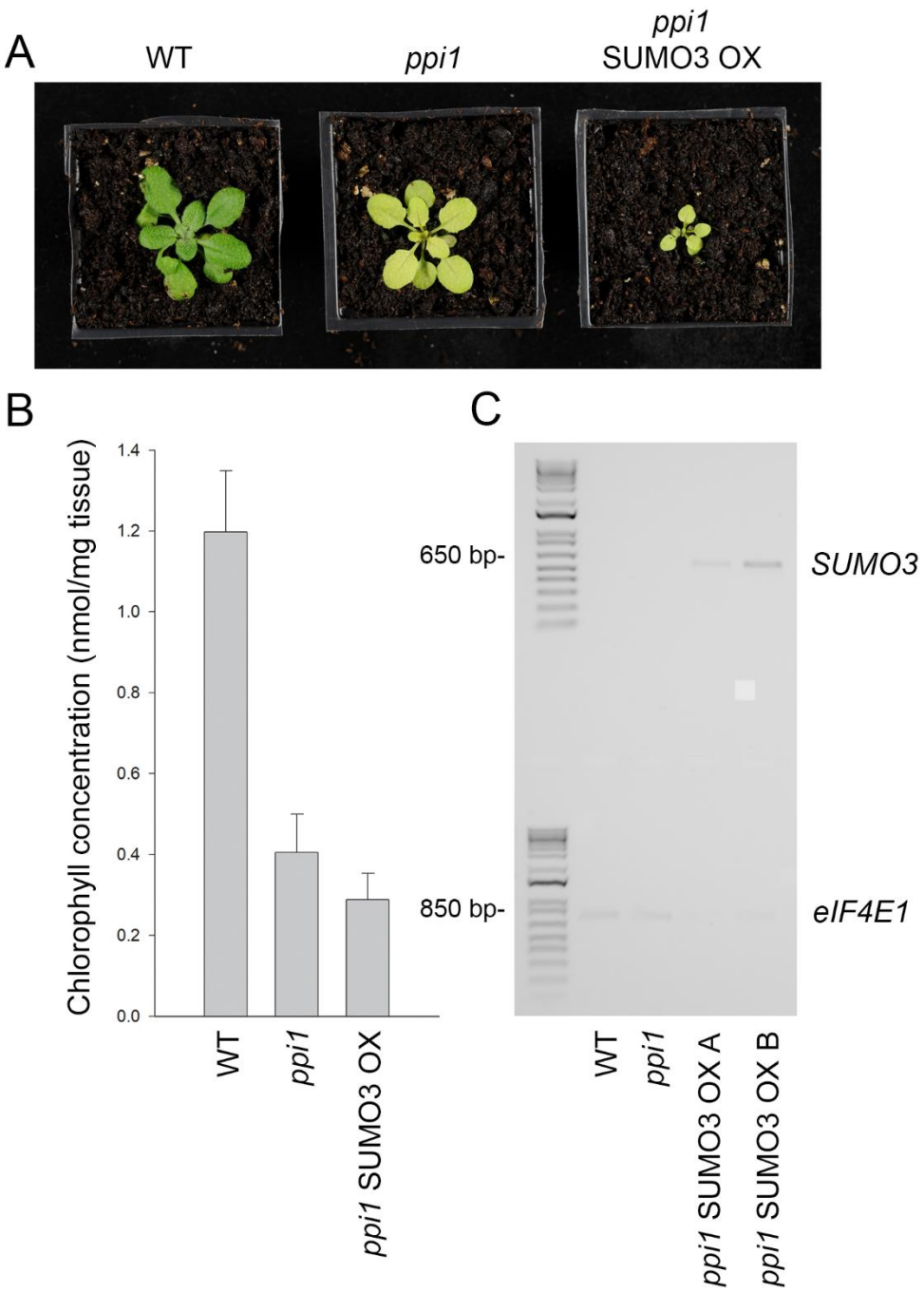


Figure 2.13

Accentuation of the *ppi1* phenotype by SUMO3 overexpression.

(A) Visual appearance of the *ppi1* SUMO3 OX transformant. Representative plants are shown. The pictured plants were grown under identical conditions for approximately four weeks.
(B) Quantification of chlorophyll accumulation in the *ppi1* SUMO3 OX transformant. Measurements were taken from rosette tissue on the day of photography (A). Error bars represent standard deviation from the mean (n = 8). There were significant differences between the *ppi1* SUMO3 OX plants and *ppi1* (Two-tailed t-test, unpaired samples, T = 2.99, p = 0.008688).

(C) Analysis of 35S:SUMO3 expression in the transformant by RT-PCR. RNA samples were taken from four-week-old rosette leave and used as a template for cDNA synthesis. Primers specific to *SUMO3* and the reference gene *eIF4E1* were used in a series of RT-PCR experiments. A limited number of amplification cycles were employed to prevent saturation (25), and the products were analysed via agarose gel electrophoresis. Amplicon size is indicated by the inclusion of a standard DNA ladder (see Chapter 6).

3 Molecular evidence that TOC proteins physically interact with the SUMO system and are SUMOylated

3.1 Abstract

The experiments described in the previous chapter demonstrated a genetic link between the SUMO system and plastid protein import. To extend this work, a series of molecular and biochemical experiments were carried out to determine whether key components of the SUMO system physically interact with TOC proteins, and whether TOC proteins are SUMOylated. The E2 SUMO conjugating enzyme SCE1 was found to interact with a range of TOC proteins in bimolecular fluorescence complementation (BiFC) and immunoprecipitation (IP) experiments. The E3 SUMO ligase SIZ1 was also found to interact with TOC proteins in a BiFC experiment; however, technical challenges prevented the completion of the complementary IP experiment. In addition, the E3 SUMO ligase HPY2 was found to interact with a range of TOC proteins in a BiFC experiment, and with Toc159 in an IP experiment. In a separate series of IP experiments, both Toc159 and Toc33 were found to physically associate with SUMO proteins in a manner strongly suggestive of SUMOylation: Tagged SUMO3 could efficiently pull down Toc159 and Toc33, whereas tagged SUMO2 and SUMO1 could pull down Toc159 with only moderate and weak efficiency, respectively.

Statement of authorship

Ms Na Li: Assisted with the BiFC cloning.

Dr Feiji Wu: Created the nYFP constructs used in the BiFC experiments.

3.2 Introduction

The previous chapter described a series of genetic experiments that strongly imply a link between the action of the SUMO system and plastid protein import. Those experiments revealed a general genetic link between the two processes, and also hinted at the identity of the specific SUMO isoforms and E3 SUMO ligases involved. Some of the genetic experiments generated slightly puzzling results: while the null mutant experiments strongly implicated SUMO1 and SUMO2 in the regulation of plastid protein import, the corresponding overexpression experiments implied a stronger role for the minor isoform SUMO3. The genetic experiments also implied a role for the E3 SUMO ligase SIZ1 in plastid protein import, but the other major E3 SUMO ligase, HPY2, was not investigated in those experiments due to a lack of suitable mutants.

To demonstrate a direct, functional link between the SUMO system and the TOC complex, a series of molecular experiments were carried out. The aims of these experiments were to demonstrate a physical interaction between key SUMO system proteins and the TOC complex, and to demonstrate that TOC proteins are SUMOylated.

3.3 Results

3.3.1. The E2 SUMO conjugating enzyme SCE1 physically interacts with TOC proteins in a BiFC experiment

The first series of experiments sought to determine whether the E2 SUMO conjugating enzyme SCE1 physically associates with TOC proteins. SCE1 is the only known E2 SUMO conjugating enzyme in *Arabidopsis* (Saracco et al., 2007), and so if the TOC complex is SUMOylated, SCE1 is highly likely to physically interact with TOC proteins. Importantly, the genetic experiments described in the previous chapter strongly implied the existence of such interactions.

The interaction between SCE1 and TOC proteins was first tested in BiFC experiments. In such experiments, genes encoding proteins that are thought to physically interact are cloned into vectors that translationally append to them sequences encoding complementary halves of a YFP protein. The two constructs are then transiently expressed (in this case, in protoplasts) and the presence of a YFP signal indicates reconstitution of functional YFP as a result of the close proximity of the two YFP halves, which implies a physical interaction between the two tagged proteins. BiFC experiments can be used to demonstrate both the presence and subcellular localisation of protein/protein interactions.

The *SCE1* CDS was cloned into the pSAT4(A)-cEYFP-N1 vector, while a number of sequences encoding TOC and negative control proteins were cloned into the complementary pSAT4-nEYFP-C1 vector. The pSAT4(A)-cEYFP-N1 vector appends the C-terminal half of YFP to the N-terminus of the expressed protein, whereas the pSAT4-nEYFP-C1 vector appends the N-terminal half of YFP to the C-terminus of the expressed protein. In a series of parallel experiments, these constructs were expressed transiently in pairs in *Arabidopsis* protoplasts. The protoplasts were subsequently assessed for YFP signals via confocal microscopy. Remarkably, SCE1 was found to physically interact with all four tested TOC proteins: Toc159, Toc132, Toc33, and Toc34 (Figure 3.1). However,

SCE1 did not robustly interact with the negative control protein ΔOEP7-cYFP. This protein comprises the transmembrane domain of plastid protein OEP7 fused to the cYFP fragment. The transmembrane domain of OEP7 is sufficient to efficiently target the full-length YFP protein to the chloroplast outer membrane (Lee et al., 2001); thus ΔOEP7-cYFP serves as a location-specific negative control. The YFP signal observed in the TOC experiments was localised to the cytosol and chloroplast envelope in all cases.

3.3.2. The E2 SUMO conjugating enzyme SCE1 physically interacts with TOC proteins in an immunoprecipitation experiment

To corroborate the findings of the BiFC experiments, a simple immunoprecipitation experiment was carried out. The *SCE1* CDS was cloned into p2GWY7, a vector that appends a YFP tag to the C-terminus of its insert (Karimi et al., 2002). The SCE1-YFP construct expressed well and showed the expected nucleocytoplasmic localisation pattern when transiently expressed in protoplasts (Figure 3.2). Next, a larger-scale protoplast transfection was carried out, and the transfected cells were solubilised and incubated with YFP-Trap magnetic beads. After incubation, the beads were isolated from the lysate via the use of a magnetic rack, then boiled in loading buffer to release bound proteins. In parallel with the SCE1-YFP construct, a construct encoding YFP-HA was used in parallel transfection and immunoprecipitation experiments. This YFP-HA construct served as a negative control. In both experiments, protein samples were collected from the following: the total lysate (i.e. the freshly solubilised protoplasts); the lysate after incubation with the YFP-Trap beads (the flow through fraction); and the YFP-Trap beads themselves (the IP fraction). Each protein sample was analysed via immunoblotting. The YFP-HA and SCE1-YFP fusion proteins both showed robust expression and strong enrichment in the IP fractions (Figure 3.2). By analysing the samples using a range of other antibodies, the SCE1-YFP fusion protein was found to be associated with native Toc159 and Toc132 but not with the negative control proteins Tic110 or Tic40. This parallels the interaction patterns exhibited by components of the CHLORAD system (Ling et al., 2012; Ling et al.,

2019). Conversely, the YFP-HA control protein could not pull down any of the tested TOC proteins, indicating the specificity of the detected SCE1 interactions. Clear evidence for interactions between SCE1-YFP and native Toc75 or Toc33 was not obtained (data not shown); however, these could be false negative results due to the significant weakness of the two relevant antibodies.

3.3.3 The E3 SUMO ligase SIZ1 physically interacts with TOC proteins in a BiFC experiment

The genetic experiments described in the previous chapter implied a role for SIZ1 in the SUMOylation of TOC proteins. To determine whether SIZ1 physically interacts with TOC proteins, BiFC experiments were carried out. The *SIZ1* CDS was cloned into the pSAT4(A)-cEYFP-N1 vector and transiently overexpressed in protoplasts alongside a series of complementary constructs encoding TOC or negative control proteins. Although the SIZ1 construct was weakly expressed relative to other constructs used in this study, a clear YFP signal was observed when the SIZ1 construct was expressed alongside constructs encoding all tested TOC proteins (Toc159, Toc132, Toc33, and Toc34) (Figure 3.3). The YFP signal observed in the TOC experiments was almost exclusively localised to the chloroplast envelope in all cases. Conversely, SIZ1 did not associate with ΔOEP7-cYFP, providing an indication of the specificity of the detected interactions. As the signals observed in the TOC experiments were comparatively weak, an additional negative control experiment was undertaken using SFR2, a protein that localises to the chloroplast outer membrane. SFR2 was not found to physically interact with SIZ1 (Figure 3.3).

3.3.4 A construct encoding the E3 SUMO ligase SIZ1 failed to express in an immunoprecipitation experiment

To corroborate the findings of the BiFC experiments, an immunoprecipitation experiment was planned. The *SIZ1* CDS was cloned into the p2GWY7 vector (used earlier in the corresponding SCE1 experiment). Unfortunately, when the resulting construct was introduced into protoplasts by transfection, it was found to be very weakly

expressed (Figure 3.4). A very small number of protoplasts showed robust fluorescence; however, the overwhelming majority did not. To guard against the possibility that there was a specific problem with the construct, the cloning was repeated and another p2GWY7 SIZ1 construct was generated. Both the old and new constructs were used in an immunoprecipitation experiment. Both SIZ1 constructs were used to co-transfect protoplasts alongside a construct encoding Toc33-HA. Unfortunately, after the samples were analysed via immunoblotting, neither SIZ1 construct was found to display detectable expression (Figure 3.4). Conversely, the Toc33-HA construct was strongly expressed, demonstrating that there were no technical problems with the transfection. In an effort to overcome this technical problem, SIZ1 could have been cloned into a different YFP-tagging vector, and the experiment could have been repeated in the presence of a proteasome inhibitor (discussed in Section 3.4). However, time constraints prevented these actions from being taken.

3.3.5 The E3 SUMO ligase HPY2 physically interacts with TOC proteins in a BiFC experiment

Alongside SIZ1, HPY2 is the other major E3 SUMO ligase described in *Arabidopsis* (Huang et al., 2009). Although HPY2 was not investigated in the genetic experiments described in the previous chapter, it could well mediate the SUMOylation of TOC proteins. HPY2 localises to the cytosol, and so it should be able to readily access the chloroplast outer membrane (Huang et al., 2009). To investigate this possibility, the *HPY2* CDS was cloned into the pSAT4(A)-cEYFP-N1 vector, and then this construct was used to co-transfect protoplasts alongside a series of TOC or negative control protein BiFC vectors. Like SCE1 and SIZ1, HPY2 was found to physically interact with all tested TOC proteins (Figure 3.5). Conversely, the HPY2 construct did not give rise to robust interaction signals when co-transfected with a vector encoding ΔOEP7-cYFP. Thus, the detected HPY2-TOC interactions were considered to be specific.

3.3.6 The E3 SUMO ligase HPY2 physically interacts with Toc159 in an immunoprecipitation experiment

To strengthen the findings of the BiFC experiments, an immunoprecipitation experiment was carried out. The *HPY2* CDS was cloned into the p2GWY7 vector, and then the resulting HPY2-YFP construct was used to transfect *Arabidopsis* protoplasts. Confocal analysis revealed strong expression of HPY2-YFP, and, as expected, the signal localised to the nucleocytoplasmic compartment (Figure 3.6). Having thus confirmed the suitability of the new construct, it was used, in parallel with a YFP-HA control construct, to transfect protoplasts for immunoprecipitation analysis as described in Section 3.3.2. Both proteins expressed well and were strongly enriched in the IP samples (Figure 3.6). Moreover, the HPY2-YFP fusion protein could pull down Toc159, but not the negative control protein Tic110. Conversely, the YFP-HA control protein could pull down neither Toc159 nor Tic110, demonstrating the specificity of the detected HPY2-Toc159 interaction.

3.3.7 Toc159 physically interacts with SUMO proteins *in vivo*

The BiFC and immunoprecipitation experiments described above demonstrated that key components of the SUMO system physically interact with TOC proteins. These physical interactions strongly imply that TOC proteins are SUMOylated; however, this was not directly demonstrated in any of the preceding experiments. To determine whether TOC proteins are SUMOylated, a further series of immunoprecipitation experiments were undertaken. The *SUMO1*, *SUMO2* and *SUMO3* CDSs were cloned into p2YGW7, a vector that appends a YFP tag to the N-terminus of its insert (Karimi et al., 2002). The p2YGW7 vector was used here instead of the p2GWY7 vector used in the other immunoprecipitation experiments as SUMO proteins undergo a C-terminal cleavage step during maturation that would result in the removal of any C-terminal tag (Wang and Dasso, 2009). All three SUMO-YFP constructs delivered robust YFP signals when introduced into protoplasts by transfection, and as expected the signals were localised in

the nucleocytoplasmic compartment (Figure 3.7). These data confirmed the suitability of the new constructs for further analysis.

The three SUMO-YFP constructs, as well as the YFP-HA control construct, were transiently expressed in protoplasts in four parallel experiments. The protoplasts were solubilised the next day, and protein samples were taken from the solubilised lysate (the total lysate fraction) and lysate-incubated YFP-Trap beads (the IP fraction). Due to space constraints on the SDS-PAGE gel, no 'flow through' samples were taken. The 'total lysate' and 'IP' samples were subsequently analysed via immunoblotting. The four experiments were resolved alongside one another on single gels. All four constructs used in this experiment expressed well and displayed strong enrichment in the IP samples (Figure 3.8). In addition, all three SUMO proteins appeared as smears on the membrane, indicating that they were conjugation competent. All three SUMO proteins could pull down native Toc159 but not the negative control protein Tic40 (Figure 3.8); SUMO3 pulled down Toc159 with strong efficiency, SUMO2 pulled down Toc159 with moderate efficiency, and SUMO1 pulled down Toc159 with weak efficiency. In all three cases, the main Toc159 band appeared as a part of a complex, high molecular weight smear. This pattern is, at least, consistent with SUMOylation. In contrast, the YFP-HA fusion protein could not robustly pull down Toc159; although a very weak single band was observed in the YFP-HA IP sample, this was considerably weaker than the corresponding single bands in the SUMO samples, and it was not accompanied by a high molecular weight-smear.

3.3.8 Toc33 physically interacts with SUMO proteins *in vivo*

The SUMOylation experiment described above produced a clear result with respect to Toc159. However, the weakness of the Toc33 and Toc75 antibodies complicated the analysis of TOC proteins other than Toc159. To circumvent this problem, the experiment was repeated under Toc33 overexpression conditions. The Toc33-HA construct used in the SIZ1 immunoprecipitation experiment (see Section 3.3.4) was used

to co-transfect the protoplasts alongside the YFP-HA and SUMO-YFP constructs. The protoplasts were solubilised and incubated with YFP-Trap beads, as in previous experiments, and then the samples were analysed via immunoblotting. The YFP-HA and SUMO-YFP constructs all expressed reasonably well and were strongly enriched in the IP samples (Figure 3.9). However, the expression was more variable than in the previous experiment, with SUMO2-YFP and SUMO3-YFP showing weaker expression than SUMO1-YFP and YFP-HA. None of the four proteins could pull down the negative control protein Tic40. Detection of Toc33 employed the Toc33 antibody, which as already noted is rather weak, and non-specifically binds to a number of proteins. Unfortunately, and unexpectedly, the antibody bound non-specifically to all three of the SUMO-YFP probes, complicating the analysis. However, the SUMO3-YFP fusion protein was clearly found to strongly and uniquely pull down Toc33 (Figure 3.9). In addition, bands of the exact expected size of Toc33 bearing one or two SUMO3-YFP moieties (75 and 114 kDa) appeared on the membrane, alongside a high molecular weight smear at the top of the membrane. This implied that Toc33 is multiply SUMOylated. Conversely, Toc33 was not pulled down by YFP-HA, SUMO1-YFP or SUMO2-YFP. However, in the case of SUMO2-YFP this could be a false negative result as the construct showed weaker expression than YFP-HA and SUMO1-YFP.

3.4 Discussion

The previous chapter described a series of experiments that demonstrated a genetic link between plastid protein import and the SUMO system. In this chapter, the mechanistic basis of this link is explored in a series of protein-protein interaction studies and SUMOylation assays.

The genetic experiments implied a role for the E2 SUMO conjugating enzyme SCE1 in TOC protein SUMOylation (Section 2.3.1). Indeed, as SCE1 is the only known E2 SUMO conjugating enzyme in *Arabidopsis* (Saracco et al., 2007), it is highly likely to physically associate with TOC proteins if the TOC complex is SUMOylated. SCE1 was therefore the ideal protein to use in the first round of protein-protein interaction studies. The BiFC experiments clearly indicated that SCE1 physically interacts with a range of TOC proteins: Toc159, Toc132, Toc33, and Toc34. Furthermore, the interactions were found to localise to the cytosol and chloroplast outer membrane. The cytosolic signal was understandable in the Toc159 and Toc132 experiments, given that at least the A-domains of both proteins are expected to be present in the cytosol (Agne et al., 2010). However, Toc33 and Toc34 are not thought to freely localise to the cytosol, and so the cytosolic signals observed in these experiments are slightly unusual. It could be that this mislocalisation is an artefact caused by the overexpression conditions (Ling et al., 2017).

The interaction between SCE1 and TOC proteins was confirmed in a simple IP experiment. SCE1-YFP was found to pull down native Toc159 and Toc132. These findings were particularly notable as SCE1 physically associates with a very broad range of substrates, and the interactions between SCE1 and its substrates are highly transient. Therefore, at any given time, only a small fraction of the pool of SCE1 will be associated with a given substrate, and this interaction is likely to be relatively weak. The fact that interactions could be captured in the IP experiments indicates that the interactions are likely to be robust.

The first set of experiments demonstrated that the genetic link between SCE1 and chloroplast protein import was linked to a physical association between SCE1 and TOC proteins. Next, experiments were performed to determine whether the E3 SUMO ligase SIZ1 also physically associates with TOC proteins. The genetic experiments described in (Section 2.3.2) hinted at such interactions. In addition, SIZ1 is the best characterised E3 SUMO ligase in *Arabidopsis*, and it is thought to physically interact with hundreds of substrates (Miura et al., 2005; Rytz et al., 2018). However, SIZ1 is also thought to be a nuclear protein, at least under standard conditions (Miura et al., 2005). Despite this, a considerable number of cytosolic proteins have been identified as SIZ1 substrates – as many as 20% of identified substrates in one analysis (Rytz et al., 2018). In addition, the original localisation study that gave rise to the view of SIZ1 as a nuclear protein did actually show very weak cytosolic localisation (Miura et al., 2005). Furthermore, in *Saccharomyces cerevisiae*, ScSiz1 (the orthologue of AtSIZ1) moves from the nucleus to the cytoplasm at certain stages of the cell cycle (Takahashi et al., 2008). It may therefore be reasonable to consider roles for SIZ1 in the SUMOylation of cytoplasmic proteins. On that basis, the SIZ1 BiFC experiments were performed. Although the SIZ1 construct used in these experiments was weakly expressed (relative to other constructs used in similar experiments), interactions between SIZ1 and all four tested TOC proteins could be observed.

The BiFC experiments suggested that SIZ1 physically interacts with TOC proteins. To corroborate this result, an IP experiment was planned. Unfortunately, two independent SIZ1-YFP constructs showed extremely weak expression in protoplasts. A very small number of protoplasts showed a robust YFP signal that localised to the nucleus and, to a lesser extent, the cytosol. This indicated that the constructs were functional but weakly expressing. It also strengthened the idea that SIZ1 can localise to the cytosol under certain circumstances. Why the SIZ1 constructs failed to express is unclear. It is noteworthy that the SIZ1 BiFC constructs also displayed relatively weak expression in

protoplasts. SIZ1 is a highly active, multifunctional protein, and so it is possible that its abundance is carefully regulated by the cell. Crosstalk between the UPS and the SUMO system is very common, and so it could be that SIZ1 is degraded in a UPS-dependent fashion under certain conditions. Many E3 ubiquitin ligases are intrinsically unstable proteins that are routinely degraded in a UPS-dependent fashion (Deshaies and Joazeiro, 2009; Ling et al., 2012; Menzies et al., 2018). This behaviour has also been observed in some E3 SUMO ligases – for example, in human cells, PIAS-type E3 SUMO ligases are targeted by SIAH2 and TRIM32, two E3 ubiquitin ligases (Depaux et al., 2007). PIAS-type proteins are thus kept at a low concentration within the cell under normal conditions, but they rapidly increase in abundance in response to certain forms of environmental stress. The IP experiment could be repeated with minor modifications to address this possibility. The protoplasts could be incubated with a proteasome inhibitor prior to collection and processing. Unfortunately, due to time constraints, this was not attempted.

My BiFC results suggested that SIZ1 physically associates with the TOC complex. However, proteins can be SUMOylated via the action of multiple E3 SUMO ligases. SIZ1 is one of two well-characterised E3 SUMO ligases in *Arabidopsis*. The other, HPY2, localises to the nucleus and cytoplasm and may therefore be in close proximity to TOC proteins (Huang et al., 2009). Indeed, in this study, HPY2 was found to physically interact with TOC proteins, both in BiFC assays and in co-IP experiments. These experiments were particularly noteworthy given the paucity of known HPY2 substrates – at time of writing, only two have been identified (Liu et al., 2016; Zhang et al., 2017).

The BiFC and immunoprecipitation experiments demonstrated that key components of the SUMO system – the E2 SUMO conjugating enzyme SCE1, and two E3 SUMO ligases – physically associate with the TOC complex. This strongly suggested that the TOC complex is SUMOylated. This notion was supported by two further experiments which demonstrated the co-immunoprecipitation of TOC proteins with different SUMO isoforms. In the first experiment, the interaction between SUMO proteins

and Toc159 was tested. SUMO3 very efficiently pulled down Toc159, while SUMO2 pulled down Toc159 with moderate efficiency. SUMO1 only weakly pulled down Toc159. These results strongly suggested that Toc159 is SUMOylated by multiple SUMO isoforms. However, there was one complicating factor. Toc159 appears as a complex ladder when resolved via immunoblotting; this is due to A-domain cleavage and complex regulation by multiple systems of post-translational modification. To provide 'gold standard' evidence of SUMOylation in IP experiments, it is necessary to show that not only can SUMO probes pull down a protein of interest, but that, after immunoblotting, a fraction of the protein of interest migrates to a higher position on the membrane. This demonstrates that the protein of interest is covalently bound to one or more SUMO moieties. A similar, protoplast-based technique was used to demonstrate that SP1 is a functional E3 ubiquitin ligase (Ling et al., 2012). In an effort to identify these high molecular weight bands, the protein samples were run on an 8% acrylamide gel for about 4 hours. Toc159 did indeed appear as a complex, high molecular weight smear in the IP samples of each SUMO experiment; however, as Toc159 also appeared as a complex smear in the total lysate control samples, these smears cannot definitively be said to represent SUMOylated species (even though this is highly likely).

The second experiment sought to determine whether SUMO proteins can physically interact with Toc33. SUMO3-YFP, but neither SUMO1-YFP nor SUMO2-YFP, could pull down Toc33-HA with strong efficiency. In addition, bands of the exact expected molecular weight of Toc33 bearing one or two SUMO3-YFP moieties (75 and 114 kDa) appeared on the membrane. These bands were accompanied by a high molecular weight smear at the top of the membrane, which is suggestive of complex, multisite or chain SUMOylation. There was, however, a complicating factor – unexpectedly, the Toc33 antibody showed binding affinity for the SUMO-YFP probes. Bands representing all three SUMO-YFP proteins clearly appeared on the membrane after immunoblotting against Toc33. Why the Toc33 antibody binds these proteins is unclear. The antibody is

promiscuous – it binds a number of non-specific proteins, including one that migrates to the position directly above Toc33. It is therefore possible that the Toc33 antibody might non-specifically bind to highly expressed proteins; however, if this were the case, it might be expected to bind the highly abundant RuBisCO proteins. Another possibility is that the YFP antibody might have contaminated the plastic tubs that were used for immunoblotting. The YFP antibody is commercially produced and is extremely strong, and so it is possible that it could persist in the washed equipment. Whatever the cause of this unusual technical problem, the experiment gave a clear result for SUMO3. Surprisingly, neither SUMO1 nor SUMO2 could pull down Toc33 in this experiment. The SUMO2 result may have been caused by weak expression and enrichment of the SUMO2-YFP construct – the construct was very weakly expressed relative to the other constructs used in the experiment.

Taken together, the evidence outlined in this chapter strongly indicates that key components of the SUMO system physically interact with the TOC complex, and that key TOC proteins are SUMOylated. In addition, the SUMOylation has a bias towards SUMO3. This is consistent with the results of the overexpression experiment outlined in Section 2.3.6, and also with the findings of a high-throughput screen for SUMO substrates (Elrouby and Coupland, 2010). This screen found that SUMO3, but not SUMO1, could physically interact with Toc159. SUMO2 was not assessed. However, the reverse genetic experiments described in Section 2.3.3 are less consistent with the SUMOylation data, principally in relation to the apparent bias towards SUMO3. The reasons behind this discrepancy are unclear. One possibility is that SUMO1 and SUMO2 can compensate for the loss of SUMO3 due to a degree of redundancy. Both of these isoforms are highly abundant relative to SUMO3 (van den Burg et al., 2010), and my data have shown that at least Toc159 can be SUMOylated by SUMO2. Many SUMO substrates are SUMOylated by SUMO1 and SUMO3 (Budhiraja et al., 2009; Elrouby and Coupland, 2010), and, as SUMO1 and SUMO2 are extremely similar (93.5% amino acid sequence identity), it is

almost certain that there is significant substrate overlap between SUMO2 and SUMO3 (Saracco et al., 2007; van den Burg et al., 2010). Thus, redundancy between SUMO1/2 and SUMO3, combined with the high-level expression of the former isoforms, might mask the effect of the *sumo3* null mutation in the *ppi1* background.

The results of this chapter and Chapter 2 demonstrate a genetic and functional link between the SUMO system, plastid protein import, and, possibly, the CHLORAD pathway. However, the precise, mechanistic details of this link are still unknown. The results are consistent with a number of models, which are explored in detail in Chapter 5.

3.5 Figures and tables

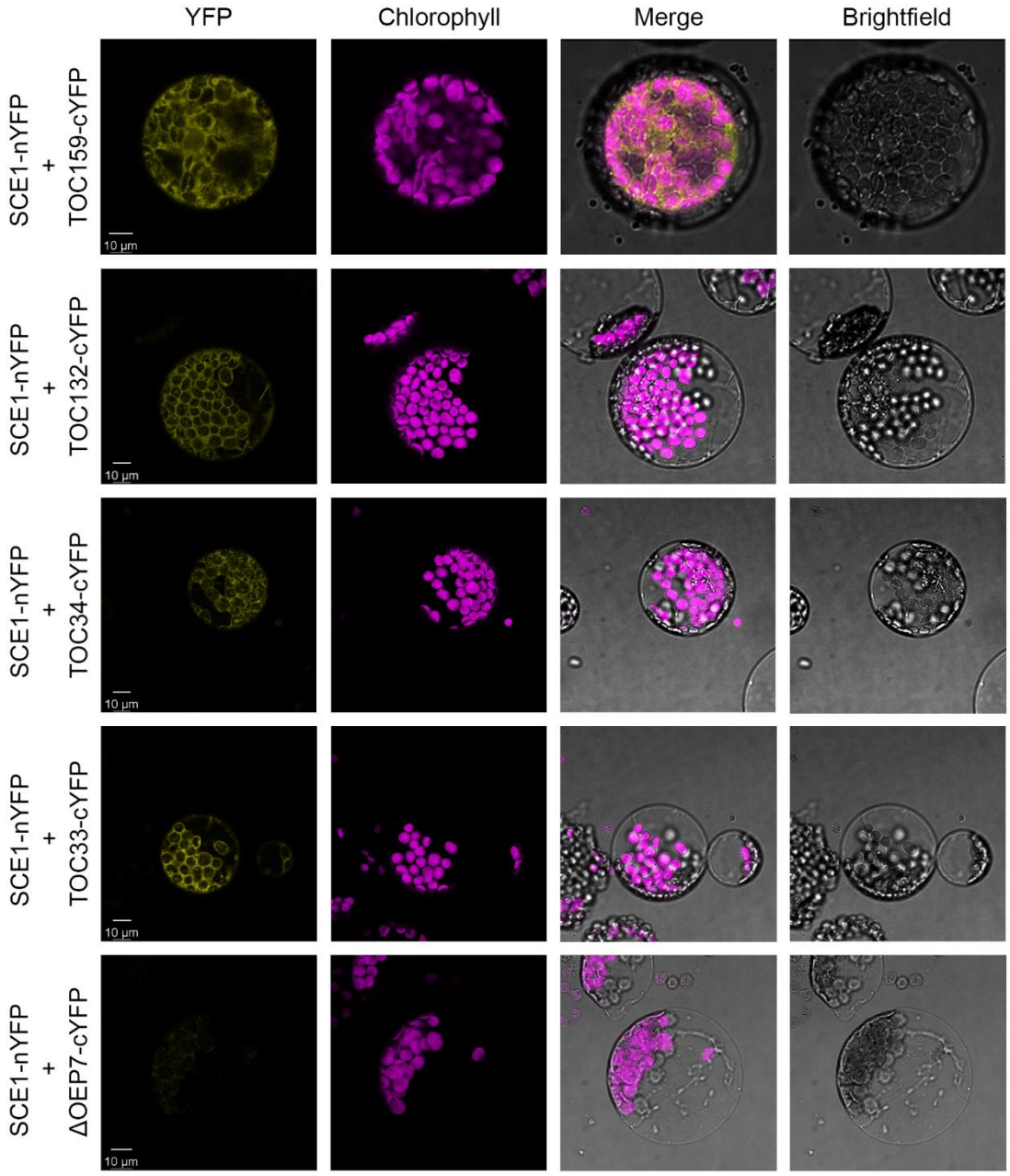


Figure 3.1

BiFC analysis of the interaction between SCE1 and TOC proteins.

The construct pairs described to the left of the figure were transiently co-expressed in protoplasts, and then the cells were assessed by confocal microscopy for reconstitution of YFP fluorescence. ΔOEP7 acted as a chloroplast-localised negative control. Representative images are shown in each case, and the exposure times and gain settings were identical in each image. nYFP and cYFP denote complementary N- and C-terminal fragments of YFP.

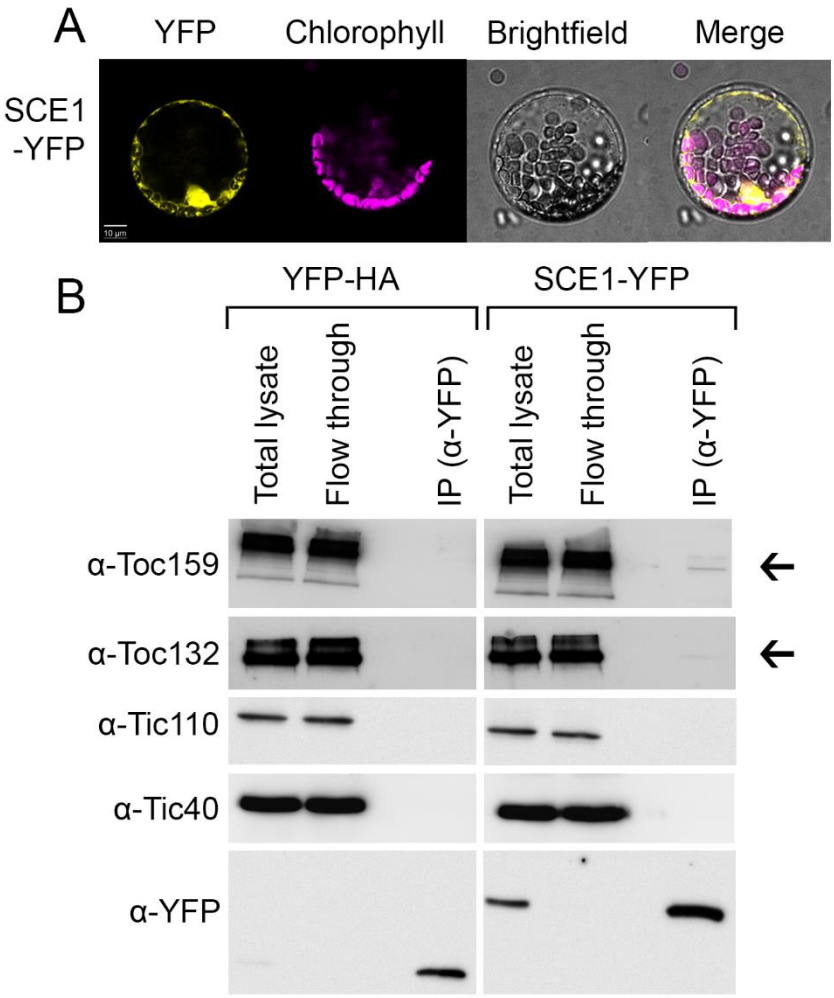


Figure 3.2

IP analysis of the interaction between SCE1 and TOC proteins.

(A) Expression analysis of the SCE1-YFP construct. The SCE1-YFP construct was transiently expressed in protoplasts and then assessed by confocal microscopy to assess both the expression and localisation of the YFP fusion protein.
(B) IP analysis of the SCE1/TOC complex interaction. The SCE1-YFP construct was transiently expressed in protoplasts as in (A), and then the ability of its product to pull down native TOC proteins was assessed by IP and immunoblotting (right hand side); parallel analyses of TIC proteins provided an indication of the specificity of the detected interactions. A construct encoding YFP-HA was similarly analysed, and this served as a negative control experiment (left hand side). IP: Immunoprecipitate. The arrows indicate the positions of the immunoprecipitated proteins.

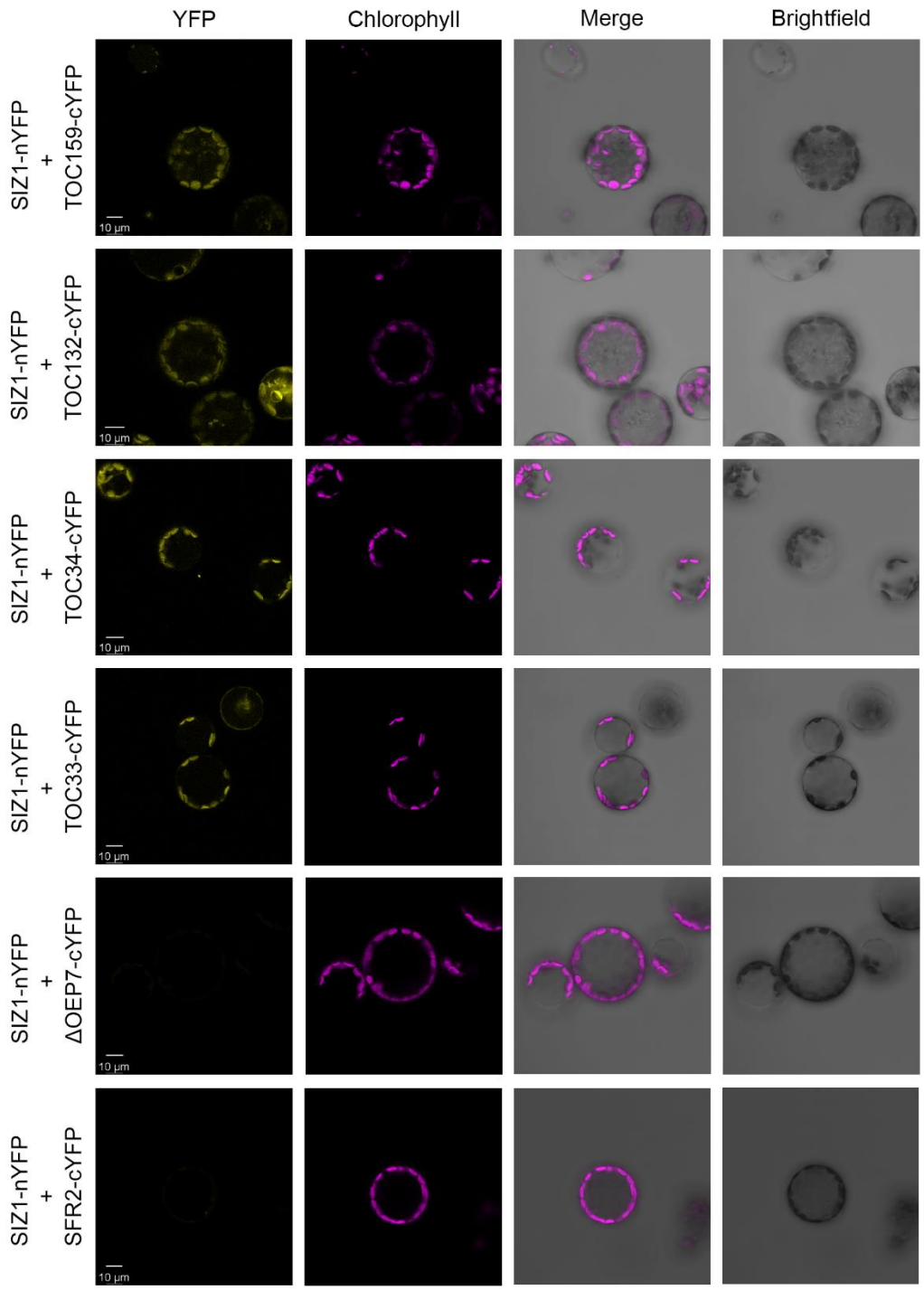


Figure 3.3

BiFC analysis of the interaction between SIZ1 and TOC proteins.

The construct pairs described to the left of the figure were transiently co-expressed in protoplasts, and then the cells were assessed by confocal microscopy for reconstitution of YFP fluorescence. ΔOEP7 and the SFR2 acted as negative controls. Representative images are shown in each case, and the exposure times and gain settings were identical in each image. nYFP and cYFP denote complementary N- and C-terminal fragments of YFP.

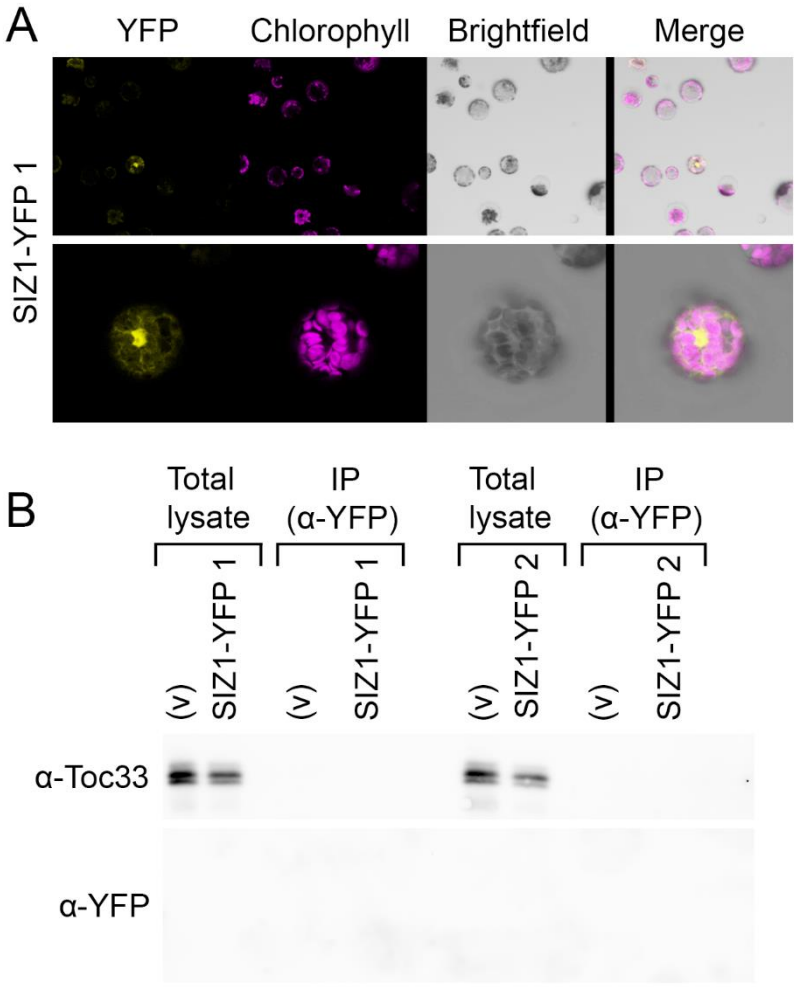


Figure 3.4

IP analysis of the interaction between SIZ1 and TOC proteins.

(A) Expression analysis of the SIZ1-YFP construct. The SIZ1-YFP construct was transiently expressed in protoplasts and assessed for both expression and localisation. Two magnifications are shown.
(B) IP analysis of the SIZ1/TOC complex interaction. Two independent SIZ1-YFP constructs were transiently expressed in protoplasts (SIZ1-YFP 1 and 2). A construct encoding Toc33-HA was co-expressed in both cases. (v): Empty vector (p2GWY7), a negative control. IP: Immunoprecipitate.

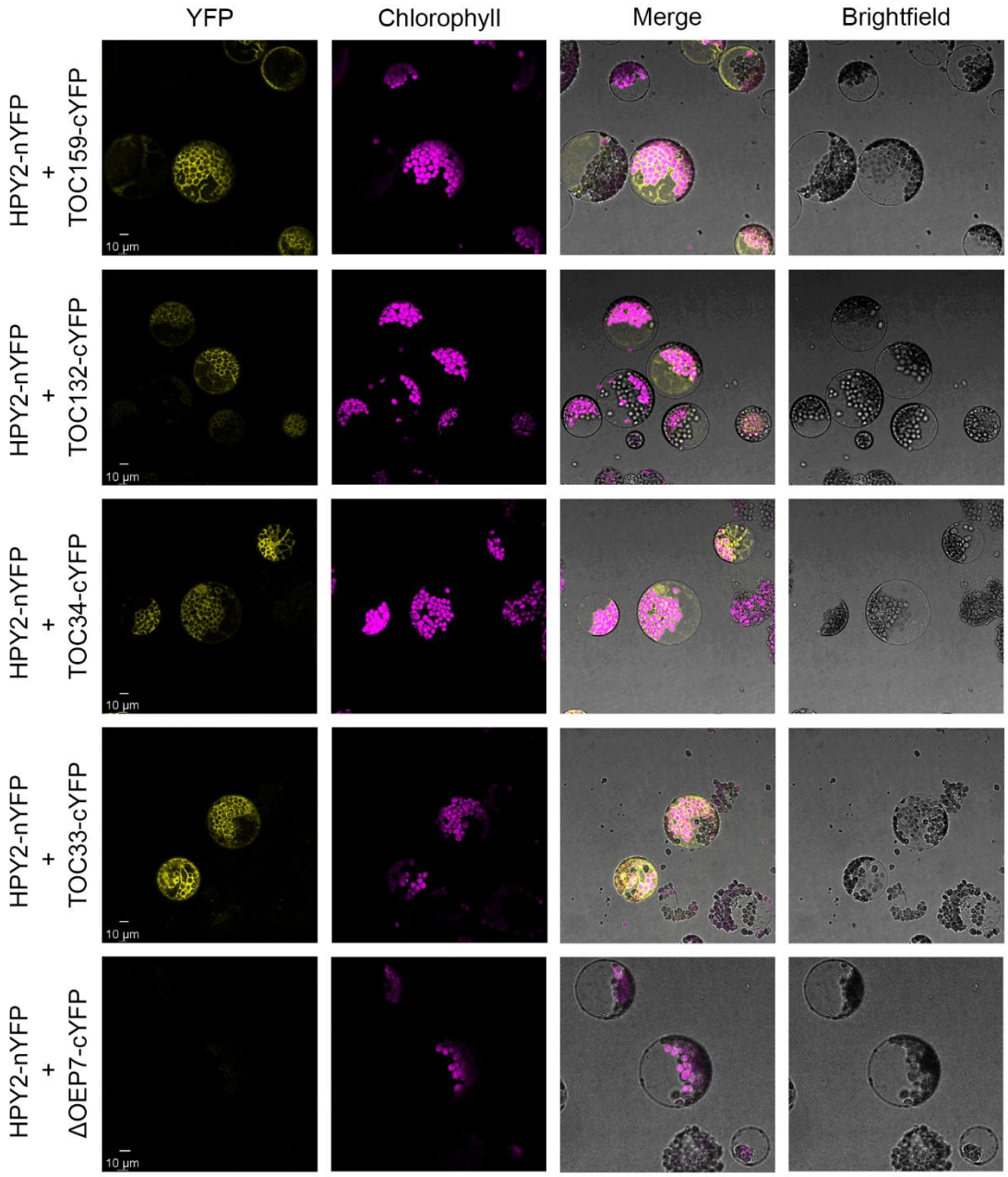


Figure 3.5

BiFC analysis of the interaction between HPY2 and TOC proteins.

The construct pairs described to the left of the figure were transiently co-expressed in protoplasts, and then the cells were assessed by confocal microscopy for reconstitution of YFP fluorescence. ΔOEP7 acted as a chloroplast-localised negative control. Representative images are shown in each case, and the exposure times and gain settings were identical in each image. nYFP and cYFP denote complementary N- and C-terminal fragments of YFP.

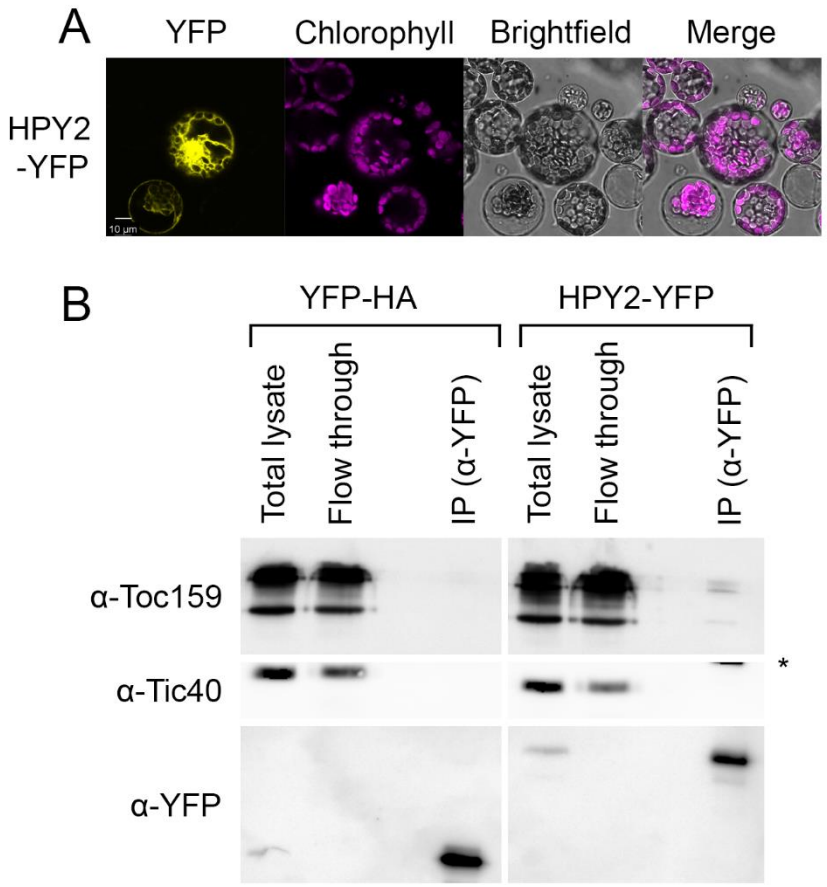


Figure 3.6

IP analysis of the interaction between HPY2 and Toc159.

(A) Expression analysis of the HPY2-YFP construct. The HPY2-YFP construct was transiently expressed in protoplasts and assessed for both expression and localisation.

(B) IP analysis of the HPY2/TOC complex interaction. The HPY2-YFP construct was transiently expressed in protoplasts as in (A), and then the ability of its product to pull down native Toc159 was assessed by IP and immunoblotting (right hand side); parallel analyses of TIC proteins provided an indication of the specificity of the detected interactions. A construct encoding YFP-HA was similarly analysed, and this served as a negative control experiment (left hand side). The asterisk indicates the position of a YFP band on the Tic40 immunoblot – the YFP membrane was reprobed with the Tic40 antibody. IP: Immunoprecipitate.

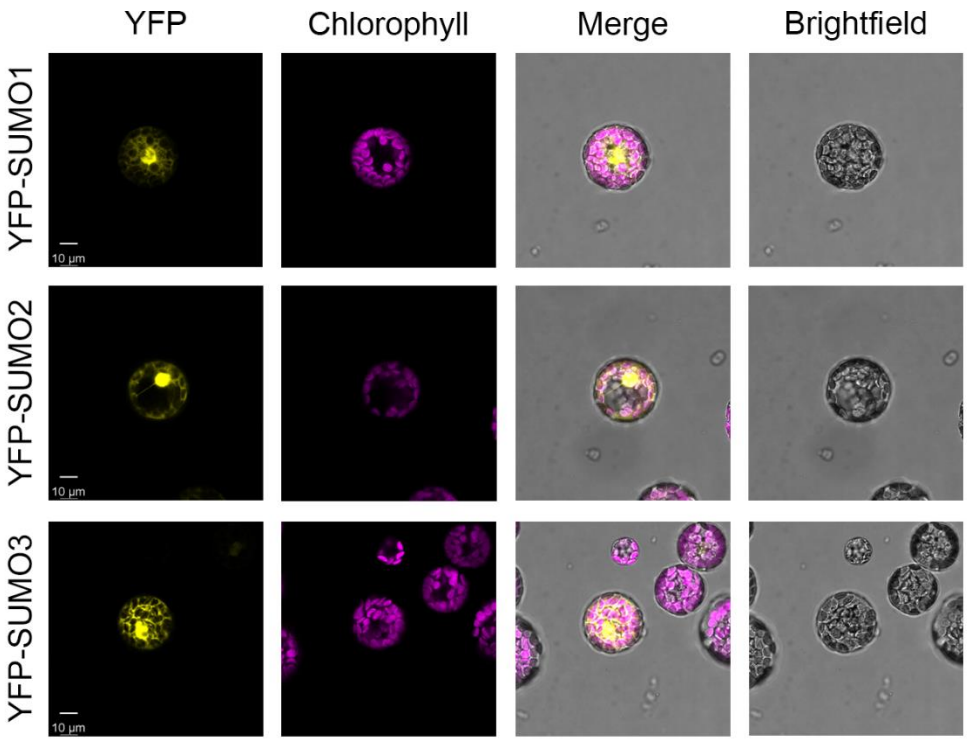


Figure 3.7

Expression analysis of the three SUMO-YFP constructs.

Expression analysis of the SUMO-YFP constructs. The SUMO-YFP constructs were transiently expressed in protoplasts and assessed for both expression and localisation.

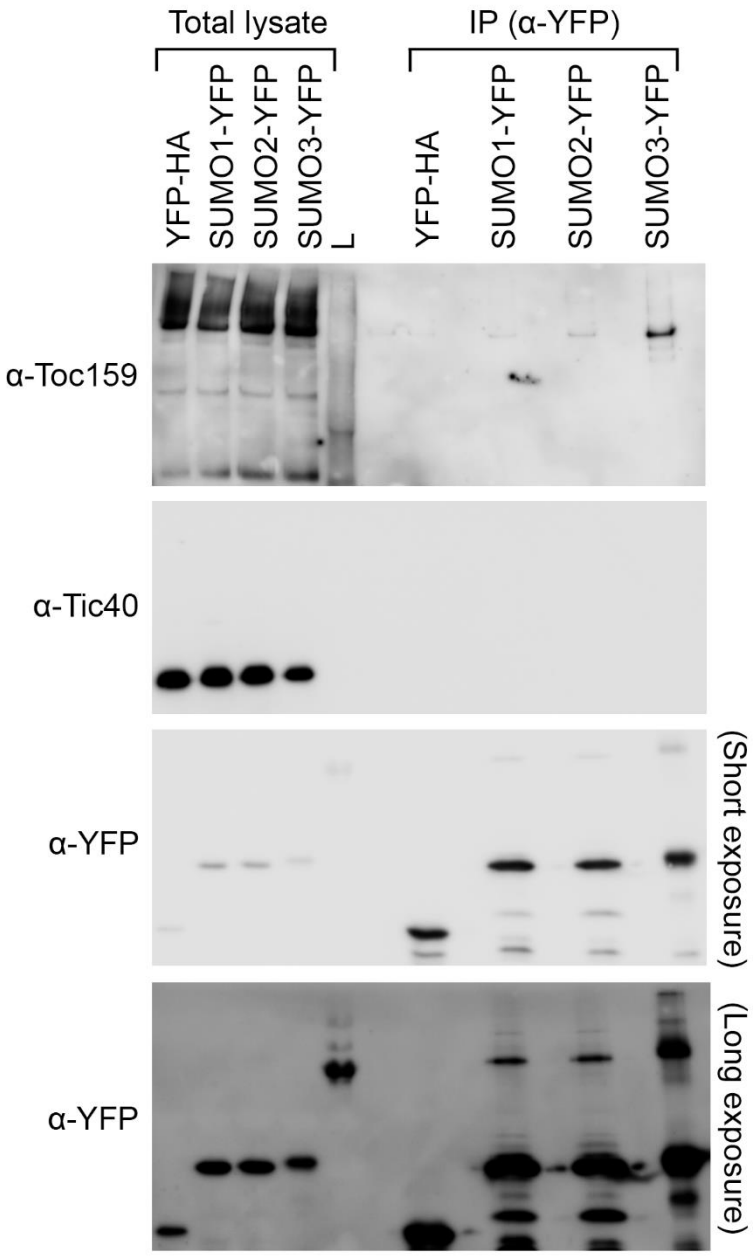


Figure 3.8

All three SUMO proteins physically interact with Toc159.

All three SUMO-YFP constructs were transiently expressed in protoplasts, and then subjected to IP analysis. YFP-HA was similarly analysed and served as a negative control. All three SUMO-YFP probes were assessed for their ability to pull down native Toc159, by analysing the IPs by immunoblotting. Tic40 served as a negative control protein. The ability of all three SUMO-YFP probes to form high-molecular weight conjugates was also assessed – the YFP immunoblot was imaged twice, once with a short exposure and once with a long exposure. In the long exposure, each SUMO protein appeared as a smear, indicating that all three probes are conjugation competent. IP: Immunoprecipitate. L: Ladder of molecular weight standards.

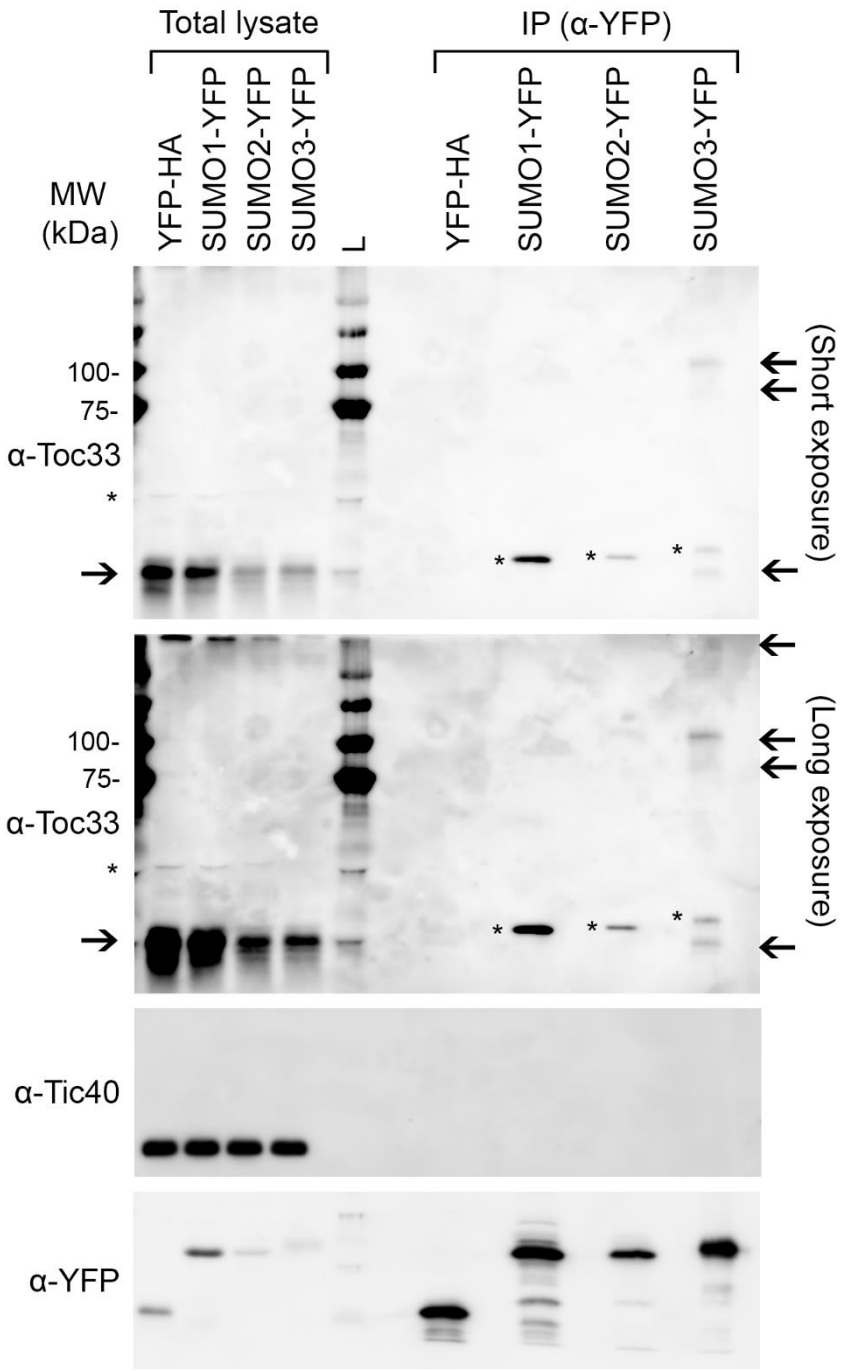


Figure 3.9

SUMO3 physically interacts with Toc33.

All three SUMO constructs were transiently expressed in protoplasts, and then subjected to IP analysis. YFP-HA was similarly analysed and served as a negative control. Toc33-HA was co-transfected in each case, to aid detection of this component. All three SUMO-YFP probes were assessed for their ability to pull down Toc33-HA, by analysing the IPs by immunoblotting. Tic40 served as a negative control protein. Arrows on the left side of the figure indicate the position of Toc33-HA. Arrows on the right of the figure indicate different forms of immunoprecipitated Toc33-HA: On the 'short exposure' panel, the

bottom arrow indicates unmodified Toc33-HA; the middle arrow indicates singly SUMOylated Toc33-HA (75 kDa); and the top arrow indicates doubly SUMOylated Toc33-HA (114 kDa). In the ‘long exposure’ panel, the same bands are highlighted, and the arrow at the top indicates high molecular weight species (possibly polySUMOylated Toc33-HA). Asterisk: non-specific bands. IP: Immunoprecipitate. L: Ladder of molecular weight standards. MW: Molecular weight standards (annotated)

4 A forward genetic screen for new regulators of the stability and function of the TOC complex

4.1 Abstract

Forward genetic screens facilitate the identification of genes that have a direct bearing on traits of interest. A recent genetic screen identified the CHLORAD pathway - a pathway for the extraction, ubiquitination, and degradation of chloroplast proteins. In this 'suppressor' screen, the pale Toc33 null mutant *ppi1* was mutagenized and plants displaying synthetic recovery phenotypes were isolated and characterised. To complement this work, a similar screen was performed. A new plastid import mutant – *fts1/ppi2-3* – was used as a starting genotype in a second order screen. The *ppi2-3* mutant bears a nonsense mutation in the *TOC159* locus (chromosome 4, c1109262t), resulting in the formation of a slightly truncated protein. Unlike *toc159* null mutants, *ppi2-3* can survive to maturity and set viable seed. The *ppi2-3* mutant was phenotypically characterised and was found to be suitable for use in a forward genetic screen. Around 29,000 *ppi2-3* seeds were chemically mutagenized, and seedlings displaying phenotypic recovery (i.e., plants that were greener than their neighbours and the *ppi2-3* progenitor line) were isolated. A total of 39 viable mutants were initially identified. Of these, 28 were found to lack obvious growth defects, such as dwarfism which indirectly cause a suppression effect. These 28 mutants were subjected to detailed phenotypic characterisation, and all were shown to retain a stable, invariant suppressor phenotype that is apparent in mature plants and is stably transmitted between generations.

4.2 Introduction

Toc159 is an indispensable component of the TOC complex and it plays a fundamental role within both chloroplast protein import and chloroplast biogenesis (Jarvis, 2008). Toc159 is an unusual and sophisticated TOC component, and its function and regulation are still incompletely understood.

Toc159 has a complex structure, comprising an N-terminal acidic domain (the 'A' domain), a GTPase domain (the 'G' domain), and a C-terminal membrane insertion domain (the 'M' domain). As outlined in detail in Chapter 2, the A-domain is readily cleaved and may be functionally important for both chloroplast protein import and substrate selection (Bolter et al., 1998; Inoue et al., 2010). Toc159's activity is regulated by a GTPase cycle, and, as such, the G-domain appears important for function (Agne et al., 2009). Toc159 is anchored in the chloroplast outer membrane via its M-domain. The M-domain, while lacking sequence features typical of a transmembrane domain, is nonetheless required for membrane insertion (Kessler et al., 1994). Interestingly, the A-domain fragment of Toc159 resides in both the cytosol and chloroplast outer membrane; however, the physiological significance of this localisation pattern is now debated (Agne et al., 2010).

Toc159 is thought to play a central role within chloroplast protein import. Although the precise mechanisms of TOC complex function are still debated, the general view is that Toc159 and Toc33 recognise incoming preproteins and guide them to the Toc75 core of the complex (Schnell, 2019). Toc159 and Toc33 both contain GTPase domains, and the two proteins are thought to function as a GTP-regulated 'switch' that controls access to the TOC complex channel.

Toc159 has a complex role within the cell, and, as such, it is regulated by a number of systems of post-translational modification. Toc159 is ubiquitinated via the action of SP1 and other currently unidentified UPS proteins (Ling et al., 2012;

Shanmugabalaji et al., 2018). As revealed in this thesis, Toc159 is also SUMOylated. In addition, Toc159 is hyperphosphorylated, possibly via the action of kinase at the outer chloroplast membrane 1 (KOC1), a kinase that localises to the chloroplast outer envelope (Agne et al., 2010; Zufferey et al., 2017).

Toc159 is a fundamentally important protein with complex and indispensable roles within the plant cell. However, Toc159's function and regulation are still incompletely understood. It is likely that aspects of Toc159's behaviour – its function, integrity, stability, and subcellular localisation - are modified via the action of currently unidentified proteins. Identifying these hypothetical partners could lead to a better understanding of Toc159's multifaceted role within the cell.

Forward genetics offers a powerful, unbiased approach towards identifying new genes that have a bearing on traits of interest. As discussed in previous sections, SP1 and SP2 were identified in a forward genetic screen for extragenic suppressors of *ppi1*, a Toc33 null mutant. A similar screen for suppressors of a pale *toc159* mutant could potentially identify interesting new regulators of Toc159; however, until recently, the albino phenotypes of all characterised *toc159* mutants made such a screen impossible.

In 2016, a new *toc159* mutant was identified (Woodson et al., 2015). The mutant was identified in a genetic screen and named ferrochelatase-two-suppressor 1 (*fts1*). The *fts1* mutant was recently renamed *ppi2-3*, and it will be referred to as such from this point on (Ling et al., 2019). The *ppi2-3* mutant carries a C-to-T mutation in the major exon of Toc159 (chromosome 4, c1109262t). Upon translation, this causes in a nonsense mutation at amino acid 1472. This premature stop codon results in a truncated product lacking the last 31 amino acids of the M-domain. As the M-domain is required for the insertion of Toc159 into the membrane, this protein is expected to show a weakened interaction with the rest of the TOC complex and compromised chloroplast localisation.

The *ppi2-3* mutant is notable in that it can survive to adulthood, flower, and set viable seed. This distinguishes it from all other characterised *toc159* mutants, which are null alleles causing seedling lethality. In the work described here, the *ppi2-3* mutant was phenotypically characterised and used as starting material in a new suppressor screen. A number of *ppi2-3* suppressor mutants were identified and characterised in detail.

4.3 Results

4.3.1 Phenotypic characterisation of *ppi2-3*

First, the new *ppi2-3* mutant was phenotypically characterised. To be suitable for use in a forward genetic screen the mutant was required to have a clear, stable phenotype that does not substantially vary between plants. Fortunately, *ppi2-3* was found to display a very clear mutant phenotype. The plants are dwarfed and extremely chlorotic: plants grown on soil for approximately 4 weeks show an 87% decrease in chlorophyll concentration relative to wild-type control plants (Figure 4.1). In addition, visual inspection of plants grown on soil or *in vitro* indicated that the degree of chlorosis is highly similar between individual *ppi2-3* plants (Figure 4.1).

4.3.2 Identification of new *ppi2-3* suppressor mutants

The phenotypic characterisation of *ppi2-3* indicated that the mutant was suitable for use in a forward genetic screen. Thus, a large volume of *ppi2-3* seed was harvested for this purpose. Approximately 29,000 *ppi2-3* seeds were chemically mutagenized in accordance with a published protocol (Leyser and Furner, 1992). Briefly, the seeds were exposed to the chemical mutagen ethyl methanesulfonate (EMS, 100 mM) for three and a half hours before being washed, collected, and sown out onto soil. The M1 seed was sown out in 100 separate batches; each batch contained approximately 270 plants grown at high density in half-trays. The plants were grown to maturity and pooled M2 seed were collected from each batch. The batches of pooled M2 seed were then grown *in vitro* and systematically screened for visible suppressor mutants (plants that appeared greener and larger than their neighbours). The presence of many albino seedlings indicated that the mutagenesis had been successful (Figure 4.2). In successful screens, 0.5-3% of M2 seedlings typically display chlorophyll accumulation defects (Leyser and Furner, 1992). Due to time constraints, only 34 of the 100 batches were screened. In total, 46 promising mutants were identified from 23 M2 batches. Each mutant was named with reference to its pool of origin (e.g. 22A) and, in cases where multiple mutants were identified in the

same pool, the order in which it was identified (e.g. 22A-1). Putative suppressor mutants were removed from plates, transferred to soil, and allowed to flower and set seed. The resulting M3 seed was then sown onto soil and the mature M3 plants were phenotypically scored. Lines that were infertile were excluded from further analysis, as were lines that did not visibly differ from the parental *ppi2-3* mutant when grown on soil. In total, 35 lines were retained for further analysis, and all of these displayed obvious phenotypic suppression. In addition, four 'borderline' putative mutant lines were retained. These either showed: complete phenotypic suppression (i.e., the plants were indistinguishable from wild-type plants), and so were possible results of contamination with wild-type seed; or only subtle phenotypic suppression.

4.3.3 Characterisation of the new *ppi2-3* suppressor mutants

The new *ppi2-3* suppressor mutants were grown on soil and phenotypically characterised with respect to chlorophyll concentration and the presence or absence of developmental defects. The plants were photographed and phenotypically scored after approximately four weeks of growth. Seven batches of plants were analysed in total, as described below.

The first batch contained six putative *ppi2-3* suppressor mutants (Figure 4.3). The magnitude of the suppression effects (i.e., the increase in leaf chlorophyll concentration relative to a single *ppi2-3* control) ranged from +57% to +167% in these mutants. Three of the lines displayed growth defects including moderately severe leaf lesions, mild dwarfism, and early flowering.

The second batch contained six putative suppressor mutants (Figure 4.4). The suppression effect magnitudes in these mutants ranged from +37% to +154%. None of the mutants in this batch showed any obvious growth defects when grown on soil.

The third batch contained six putative suppressor mutants (Figure 4.5). Suppression effect magnitudes ranged from +47% to +84%. Unfortunately, all the mutants

in this batch displayed possible developmental defects. These defects related to plant size (dwarfism and possible gigantism), phenotypic variability, and the presence of leaf lesions.

The fourth batch contained six putative suppressor mutants (Figure 4.6). Suppression effect magnitudes ranged from +60% to +309%. Three of the mutants in this batch displayed obvious or possible developmental defects. These defects related to leaf edge morphology, rosette size, and the presence of leaf lesions.

The fifth batch contained six putative suppressor mutants (Figure 4.7). Suppression effect magnitudes were highly variable between the different mutants, ranging from +73% to +324%. One of the mutants displayed a subtle leaf morphology defect, and another was mildly dwarfed.

The sixth batch contained five putative suppressor mutants (Figure 4.8). The suppression effects were moderate in this case, ranging from +45% to +80%. None of the mutants displayed any obvious developmental defect.

The seventh and final batch contained lines that were considered borderline (Figure 4.9). The lines were considered borderline for various reasons: one of the lines displayed extremely subtle phenotypic suppression, two displayed absolute phenotypic suppression, and one was contaminated with wild-type seed. The lines that displayed complete phenotypic suppression are probably descended from contaminating wild-type seed. However, until the lines have been robustly genotyped with respect to the *ppi2-3* mutation, this cannot be confirmed. The contaminated line was grown up on soil and plants displaying a *ppi2-3* phenotype were scored with respect to chlorophyll concentration. A moderate suppression effect was observed (+89%). In addition, the mutant displaying very subtle suppression was phenotypically characterised and was shown to display a mild suppression effect (+12%).

The full catalogue of putative *ppi2-3* suppressor mutants is outlined in (Table 4.1). Mutants displaying obvious growth defects were, in most cases, considered unsuitable for downstream study. The ideal mutant is one in which the suppression effect is linked to a mutation in a gene with selective and specific action linked to TOC function. The suppression effects observed in mutants with broad phenotypic defects are more likely to be secondary effects than effects that have a direct bearing on the stability and function of Toc159. After the removal of such lines, 28 mutants were left (Table 4.2). The mutants were renamed suppressor of *ppi2-3* 1 to 28 (*sop1-sop28*). These lines should be retained for future analysis.

4.4 Discussion

Suppressor screens represent a powerful and unbiased approach towards dissecting complex pathways and regulatory networks. The application of suppressor screens to chloroplast protein import recently revealed the existence of a remarkable new pathway that mediates the extraction, ubiquitination, and degradation of chloroplast proteins (Ling et al., 2012; Ling et al., 2019). This screen used the Toc33 null mutant *ppi1* as a starting genotype. In the work described in this chapter, a new mutant, *ppi2-3*, was used as a starting genotype in a similar suppressor screen. The *ppi2-3* mutant carries a mutation in the major exon of the Toc159 locus. Upon translation, this mutation induces a premature stop codon that causes a small truncation in the C-terminal membrane insertion domain of Toc159. This truncation is likely to reduce, but not abolish, the ability of Toc159 to associate with the chloroplast outer membrane. Accordingly, the *ppi2-3* mutant is small and chlorotic but viable (in contrast with *toc159* null mutants *ppi2-1* and *ppi2-2* which are albino, seedling lethal, and infertile (Bauer et al., 2000; Kubis et al., 2004)).

The first step in this analysis was the phenotypic characterisation of the *ppi2-3* mutant. The plants were found to display severe chlorosis when grown on soil and *in vitro*, and in both cases the severity of the phenotype was similar between individual plants (Figure 4.1).

Approximately 29,000 *ppi2-3* seeds were exposed to EMS, a chemical mutagen, and 46 putative suppressor mutants were initially identified in the M2 generation. Of these, 39 retained a visible suppressor phenotype when transferred to soil and were capable of setting viable seed. All 39 mutants were grown up on soil alongside relevant control plants and phenotypically characterised.

The new suppressor mutants are highly variable, with some mutants displaying a subtle suppression effect (e.g. the 3B-1 line which showed a 37% increase in chlorophyll

concentration) and others displaying very robust phenotypic suppression (e.g. the 15A-2 line which showed a 324% increase in chlorophyll concentration). A number of the lines had growth and morphology defects, and when these were excluded, 28 mutant lines were retained and prioritized for future analysis.

4.4.1 Future directions

In the next phase of this project, an immediate priority will be to backcross each of the 28 'good' mutants to the *ppi2-3* progenitor mutant, for several reasons: to conduct basic genetic analyses of inheritance and dominance; to remove additional unwanted background mutations; and to generate mapping populations for sequencing-based gene identification analysis. The resulting F1 and F2 plants should be phenotypically scored to identify single, segregating loci responsible for the observed suppression effects. Given that almost all the mutants show limited phenotypic variability, most lines are expected to carry a single causative mutation. This backcross analysis will also enable the nature of each mutation (i.e., whether it is dominant or recessive) to be determined. At time of writing, about half of the mutants have been backcrossed.

Before the causative loci are mapped, the *SP1* and *SP2* genes of each mutant will be sequenced (to determine whether they have been mutated) and analysed by RT-PCT (to determine whether their expression has been altered). The *sp1* and *sp2* mutations can suppress the plastid import mutants *ppi1* and *toc75-III-3*, and so it is likely that *sp1* and *sp2* mutants can also suppress *ppi2-3* (Ling et al., 2012; Ling et al., 2019). Any new *sp1* and *sp2* alleles will be retained but excluded from downstream analyses, as genes encoding novel participants in TOC regulation will be prioritized. Genes encoding SUMO proteins will not be similarly analysed, as mutations in most of these loci will cause severe growth defects, and mutants displaying such defects have been excluded from the analysis.

After any new *sp1* and *sp2* mutants have been removed, the catalogue of remaining mutants will be brought forward for genetic mapping. A mapping-by-sequencing

approach will be employed (Schneeberger et al., 2009). The segregating F2 populations resulting from backcrosses to the *ppi2-3* progenitor mutant will be grown up and the plants will be phenotypically scored (for the presence/absence of the suppressor phenotype). DNA will be extracted from pooled seedlings displaying the suppressor phenotype (at least 50 seedlings will be used in each extraction). The entire complement of DNA contained within each sample will be sequenced via a next-generation sequencing platform, and the resulting data will be analysed using the SHOREmap software package (Schneeberger et al., 2009). This method is a form of bulk segregant analysis and allows the identification of single nucleotide polymorphisms (SNPs) that are tightly linked to the mutant phenotype. The SHOREmap analysis will identify the narrow genomic neighbourhood of each suppressor mutation. Candidate genes contained within the genomic window will be identified, and T-DNA mutants representing all likely candidates will be obtained and crossed with *ppi2-3*. Only the T-DNA mutants affecting the causative locus are likely to induce phenotypic suppression in the *ppi2-3* background.

4.4.2 Possible suppressor loci

The suppressor mutants identified in this study could have lesions in genes that modify many aspects of Toc159's function and stability, or the function and stability of the entire TOC complex. This screen is likely to identify genes other than those identified in the screen in which CHLORAD was discovered. The first screen identified suppressors of the pale, Toc33 null mutant *ppi1*. Toc159 is much larger than Toc33, it has a more complex structure, and it seems to be subjected to a broader range of post-translational modifications (Agne et al., 2010). In addition, *ppi2-3* is a hypomorphic allele whereas *ppi1* is a null allele (Jarvis et al., 1998; Ling et al., 2019).

Toc159 is targeted by the ubiquitin-proteasome system via SP1 and, probably, by additional, currently unknown cytosolic and/or nuclear E3 ubiquitin ligases (Ling et al., 2012; Shanmugabalaji et al., 2018). These unknown E3 ubiquitin ligases could be identified in this screen alongside other UPS components (e.g. E2 ubiquitin conjugating

enzymes) and accessory proteins involved in the regulation of TOC complex stability. Toc159 (and other TOC proteins) are also regulated by phosphorylation, and so new kinases could be identified in this screen (Agne et al., 2010). In addition, new components of the CHLORAD pathway may be uncovered.

Toc159 is an unusual TOC protein in that it is present in a number of subcellular locations. It has been known for some time that the Toc159 A-domain is readily cleaved and localises to the cytosol; however, the physiological significance of this process is currently unknown (Bolter et al., 1998; Agne et al., 2010). Genes that are involved in this process could be uncovered in this screen. The discovery of such genes could provide new insights into the basic mechanisms of TOC complex function.

A recent study found that Toc159 stability is regulated by the plant hormone gibberellic acid (GA) during germination (Shanmugabalaji et al., 2018). This regulation is mediated by physical interactions between Toc159 and DELLA proteins. DELLA proteins are negative regulators of GA signalling and are degraded in a UPS-dependent fashion in response to GA accumulation (Daviere and Achard, 2013). This degradation is promoted by GA-activated GID1, the GA receptor. Prior to germination, GA levels are low and DELLA proteins bind Toc159 and promote its ubiquitination and degradation. During germination, GA levels increase, and the DELLA proteins bound to Toc159 are themselves ubiquitinated and degraded. The degradation of DELLA proteins allows Toc159 to accumulate and associate with other TOC proteins. As GA promotes Toc159 stability, genes encoding negative regulators of GA signalling may appear in the catalogue of new suppressor mutants.

4.5 Figures and tables

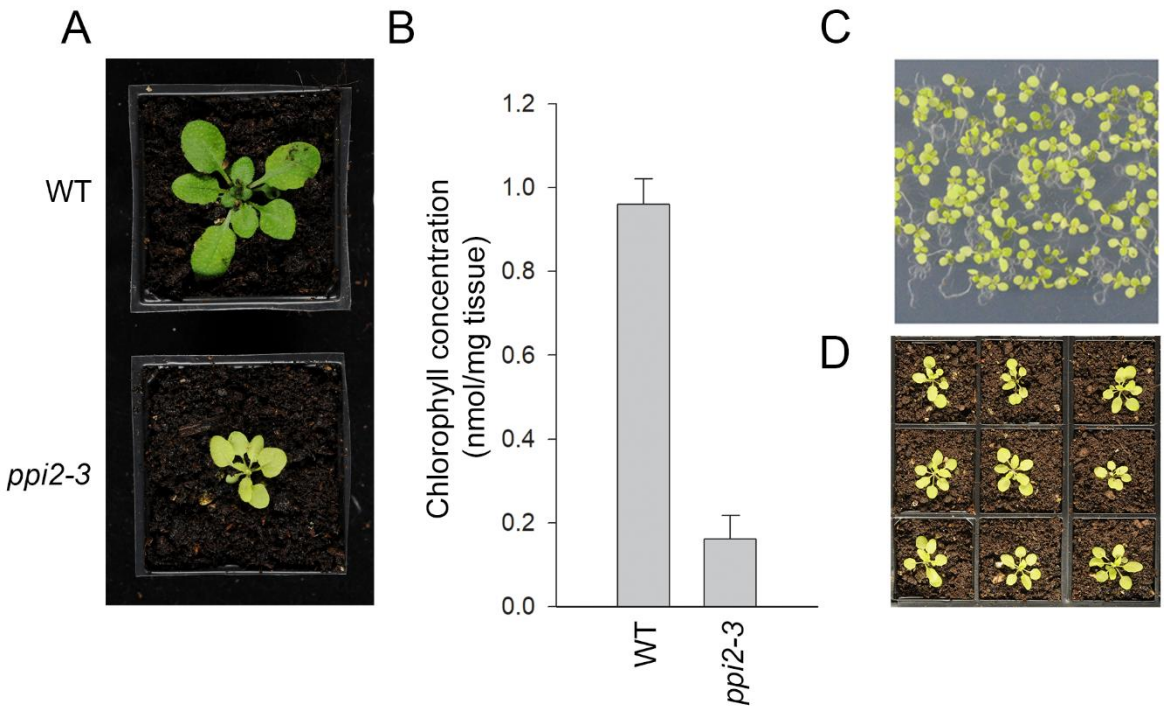


Figure 4.1

Phenotypic characterisation of the *fts1*/*ppi2-3* mutant.

(A) Visual appearance of the *ppi2-3* mutant. Representative plants are shown. Plants were grown under identical conditions for approximately four weeks.
(B) Quantification of chlorophyll concentration in *ppi2-3*. Measurements were taken from rosette tissue on the day of photography (A) using a SPAD meter. Error bars indicate standard deviation from the mean (n = 5).
(C) Analysis of phenotypic variability in *ppi2-3* when grown on MS media. The *ppi2-3* seeds were sown onto MS media and grown in a controlled growth chamber for two weeks.
(D) Analysis of phenotypic variability in *ppi2-3* when grown on soil. The pictured plants were grown under identical conditions for approximately four weeks.

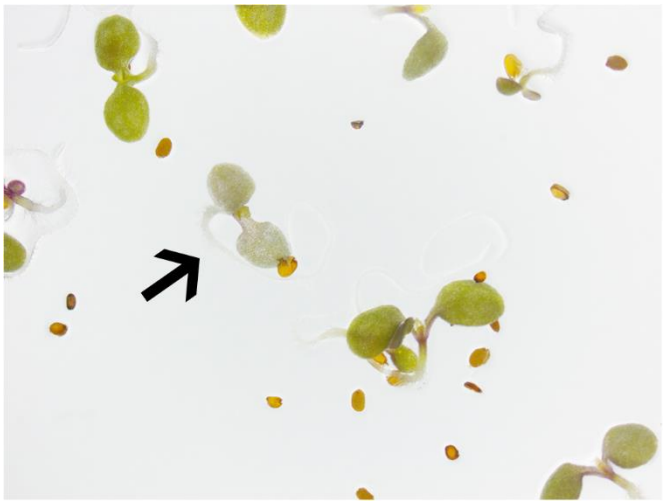


Figure 4.2

Albino seedlings are present in the mutagenized *ppi2-3* seed.

The M2 generation mutagenized seeds were sown on MS media, and the plants were grown for 2 weeks. The arrow indicates an albino seedling; these were frequently found in the pool of mutagenized seedlings. In successful screens, 0.5-3% of M2 progeny plants display chlorophyll defects (Leyser and Furner, 1992).

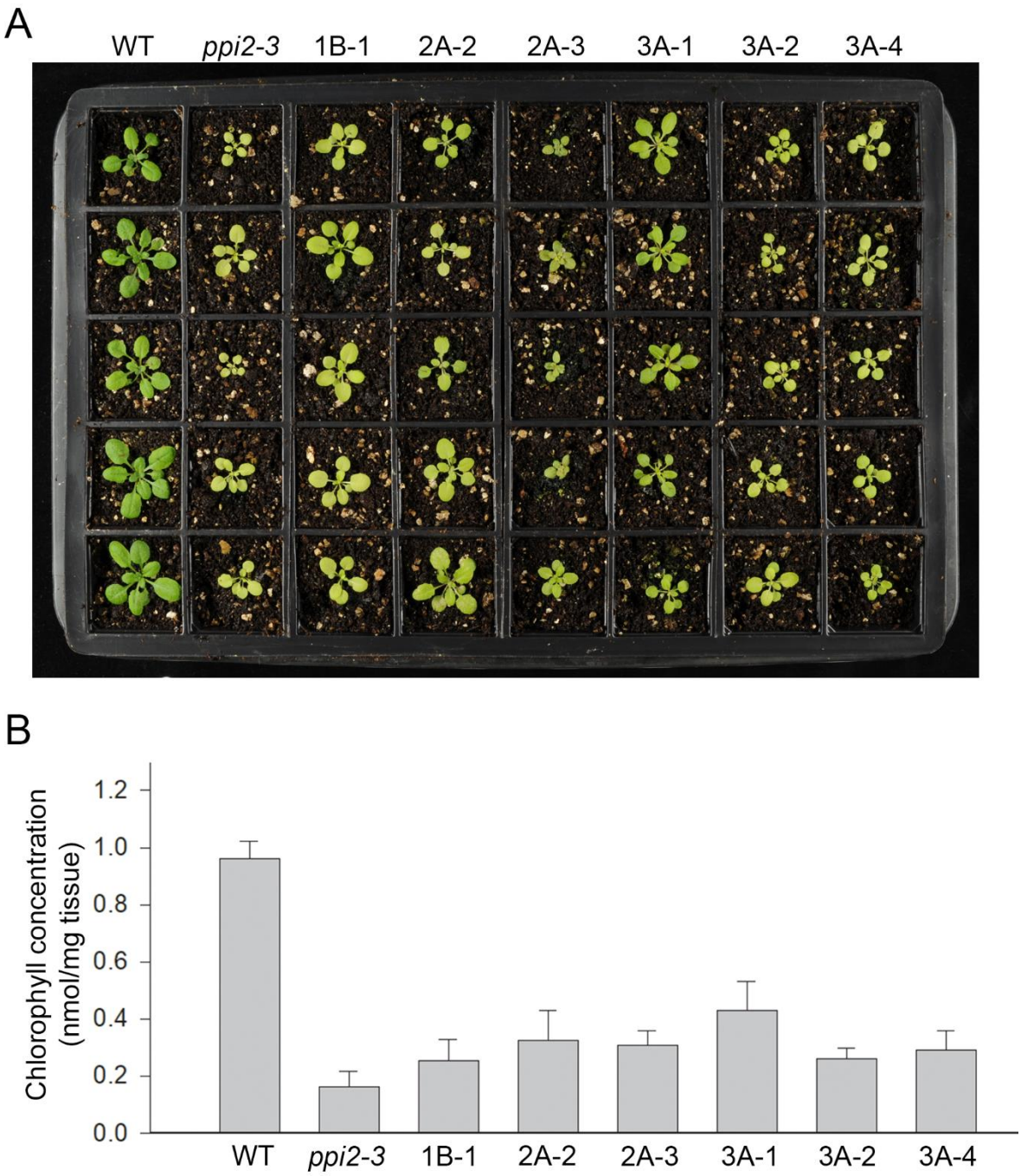


Figure 4.3

The first batch of new *ppi2-3* suppressor mutants.

(A) Visual appearance of the new *ppi2-3* suppressor mutants. The pictured plants were grown under identical conditions for 4 weeks.
(B) Quantification of chlorophyll accumulation in the new *ppi2-3* suppressor mutants. Measurements were taken on the day of photography (A) using a SPAD meter. Five plants were analysed per genotype. Error bars indicate standard deviation from the mean (*n* = 3-5).

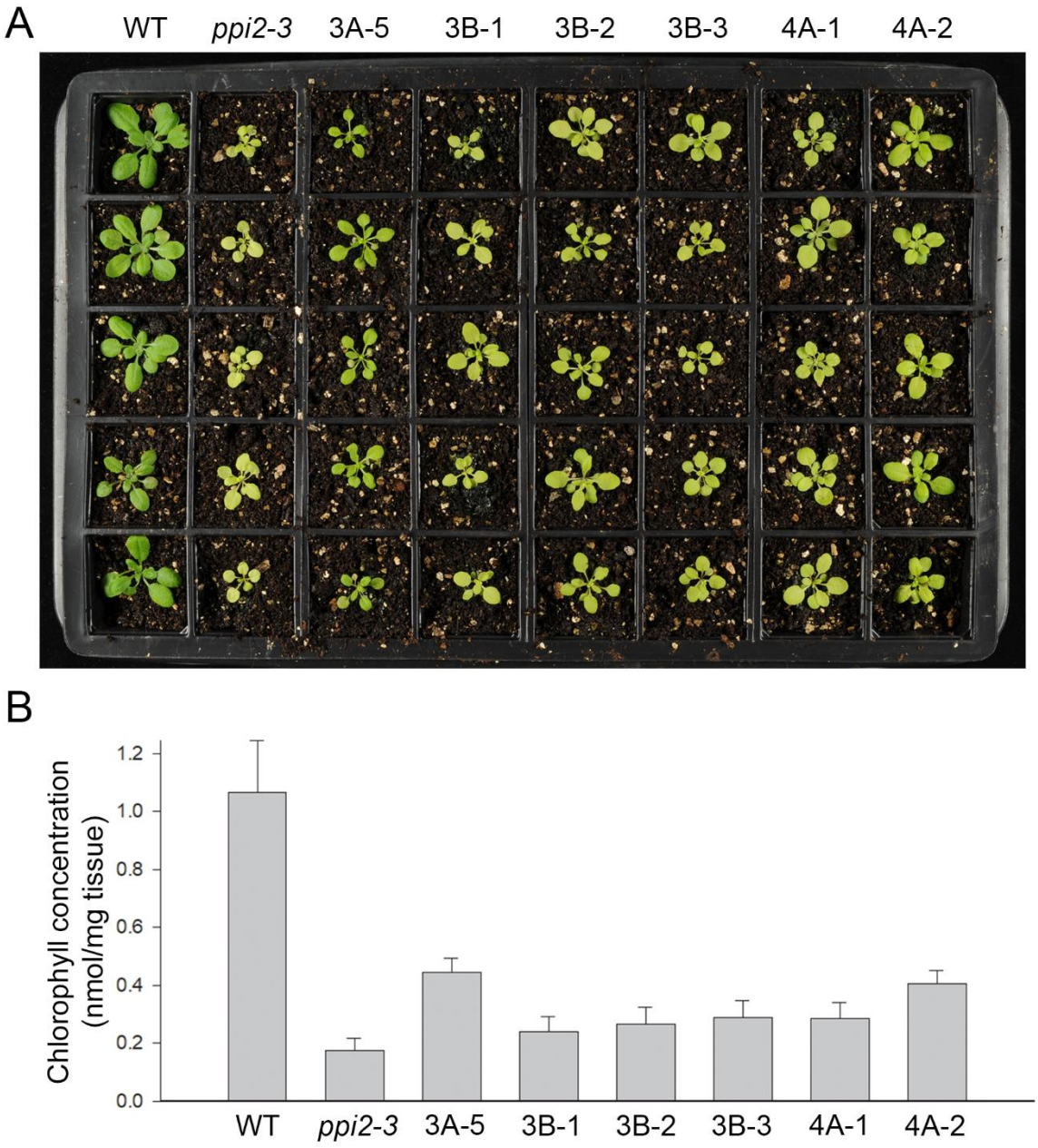


Figure 4.4

The second batch of new *ppi2-3* suppressor mutants.

(A) Visual appearance of the new *ppi2-3* suppressor mutants. The pictured plants were grown under identical conditions for 4 weeks.
(B) Quantification of chlorophyll accumulation in the new *ppi2-3* suppressor mutants. Measurements were taken on the day of photography (A) using a SPAD meter. Five plants were analysed per genotype. Error bars indicate standard deviation from the mean (*n* = 4-5).

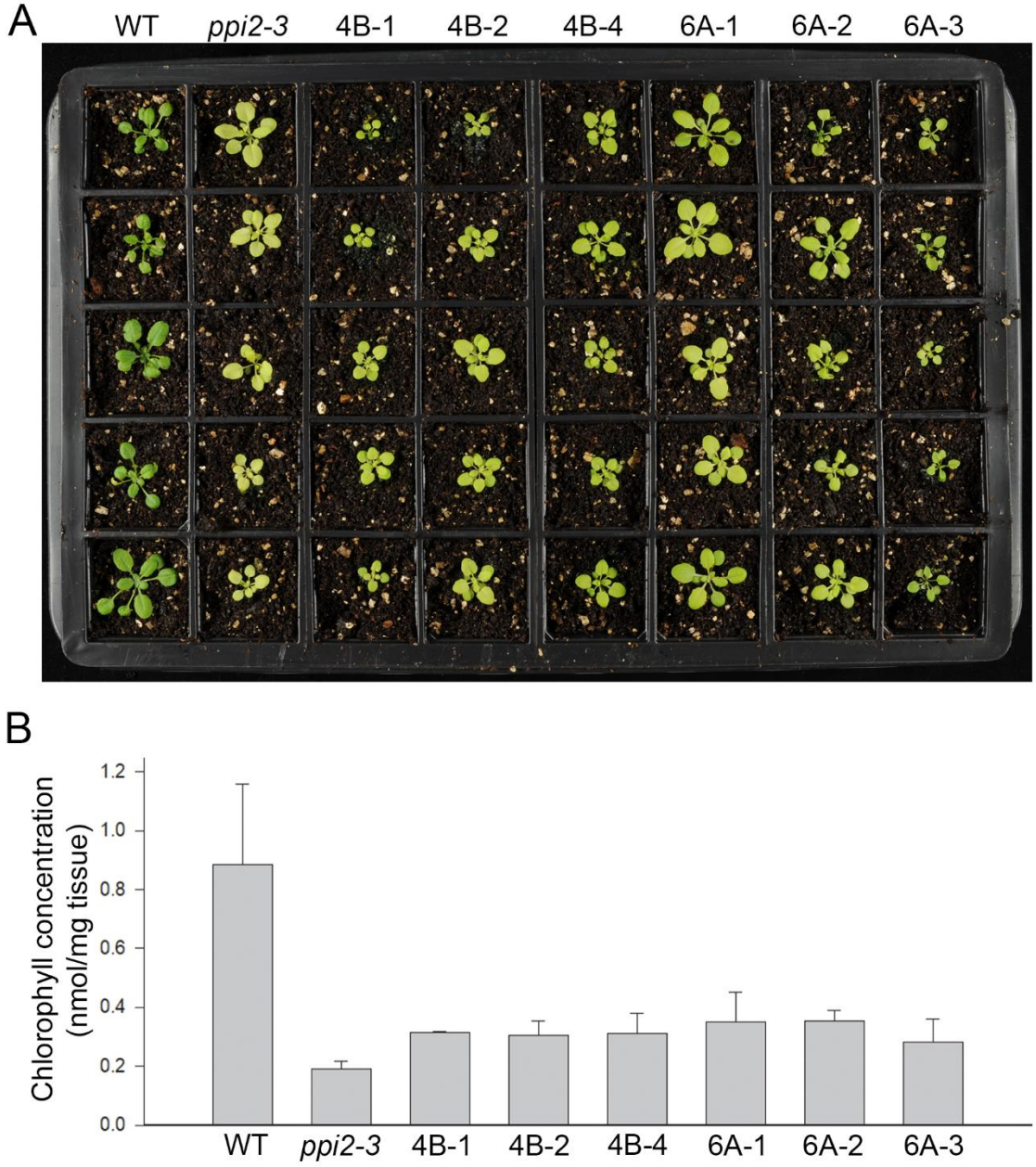


Figure 4.5

The third batch of new *ppi2-3* suppressor mutants.

(A) Visual appearance of the new *ppi2-3* suppressor mutants. The pictured plants were grown under identical conditions for 4 weeks.
(B) Quantification of chlorophyll accumulation in the new *ppi2-3* suppressor mutants. Measurements were taken on the day of photography (A) using a SPAD meter. Five plants were analysed per genotype. Error bars indicate standard deviation from the mean (n = 2-5).

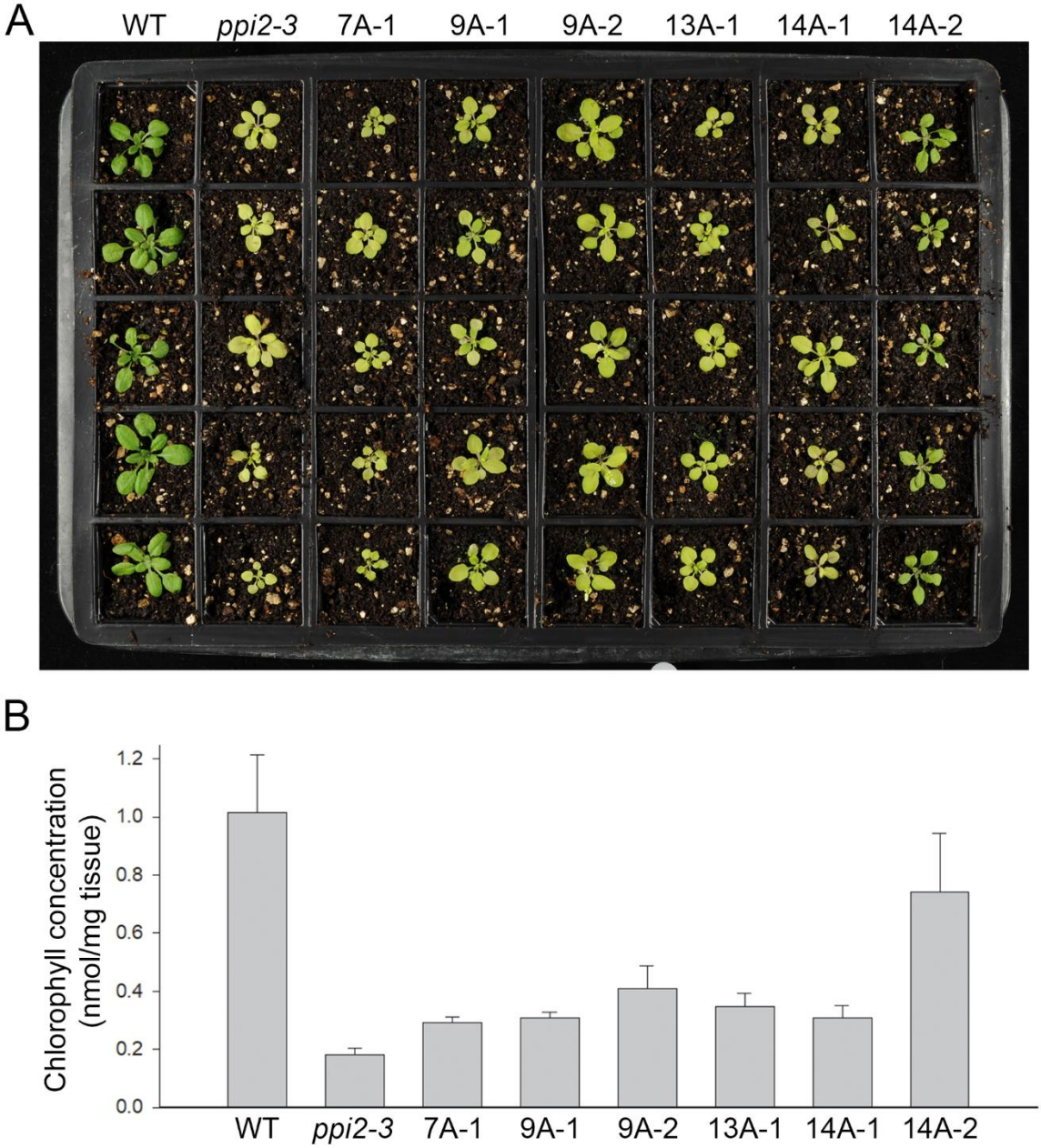


Figure 4.6

The fourth batch of new *ppi2-3* suppressor mutants.

(A) Visual appearance of the new *ppi2-3* suppressor mutants. The pictured plants were grown under identical conditions for 4 weeks.
(B) Quantification of chlorophyll accumulation in the new *ppi2-3* suppressor mutants. Measurements were taken on the day of photography (A) using a SPAD meter. Five plants were analysed per genotype. Error bars indicate standard deviation from the mean (*n* = 3-5).

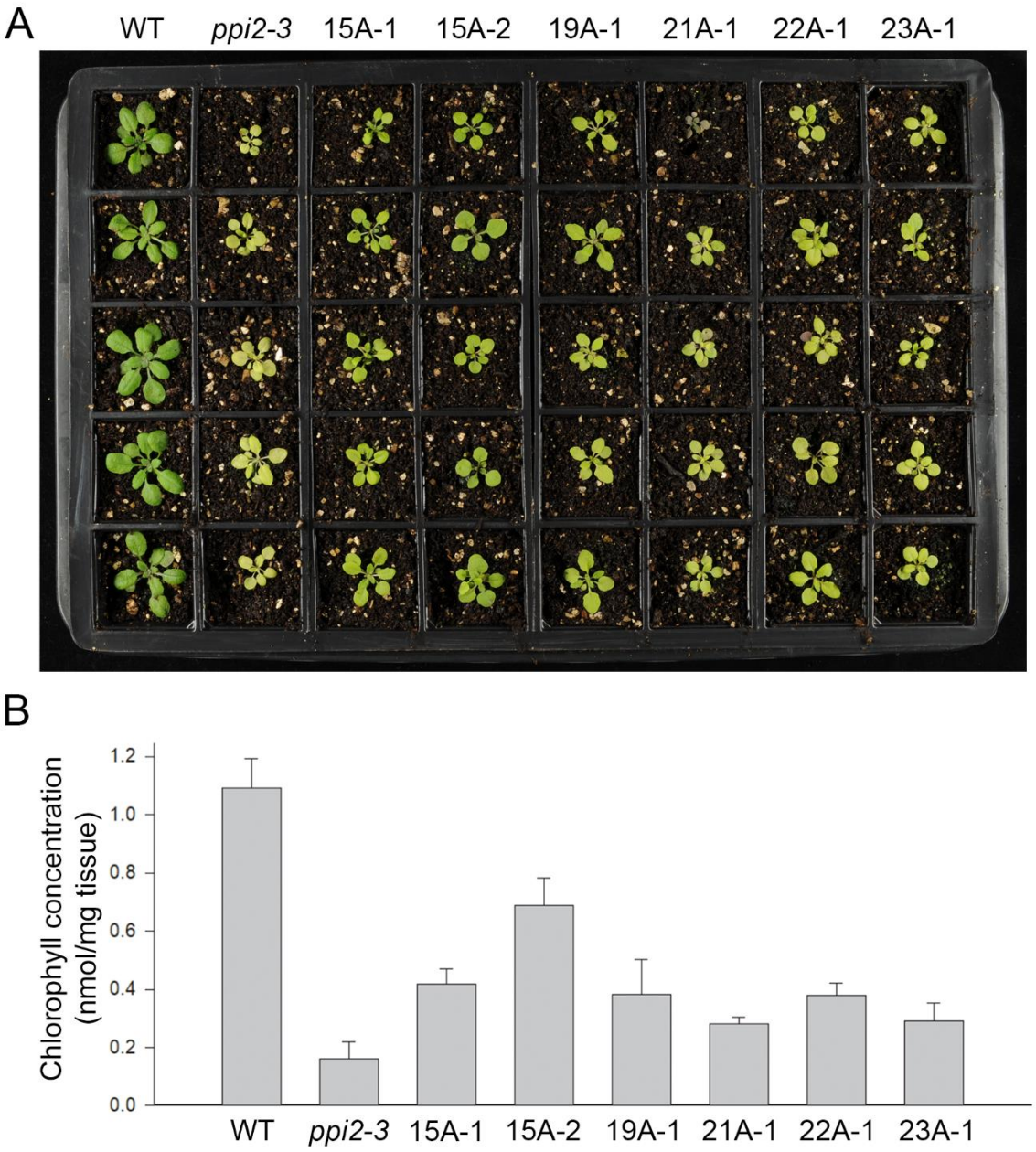


Figure 4.7

The fifth batch of new *ppi2-3* suppressor mutants.

(A) Visual appearance of the new *ppi2-3* suppressor mutants. The pictured plants were grown under identical conditions for 4 weeks.
(B) Quantification of chlorophyll accumulation in the new *ppi2-3* suppressor mutants. Measurements were taken on the day of photography (A) using a SPAD meter. Five plants were analysed per genotype. Error bars indicate standard deviation from the mean (*n* = 3-5).

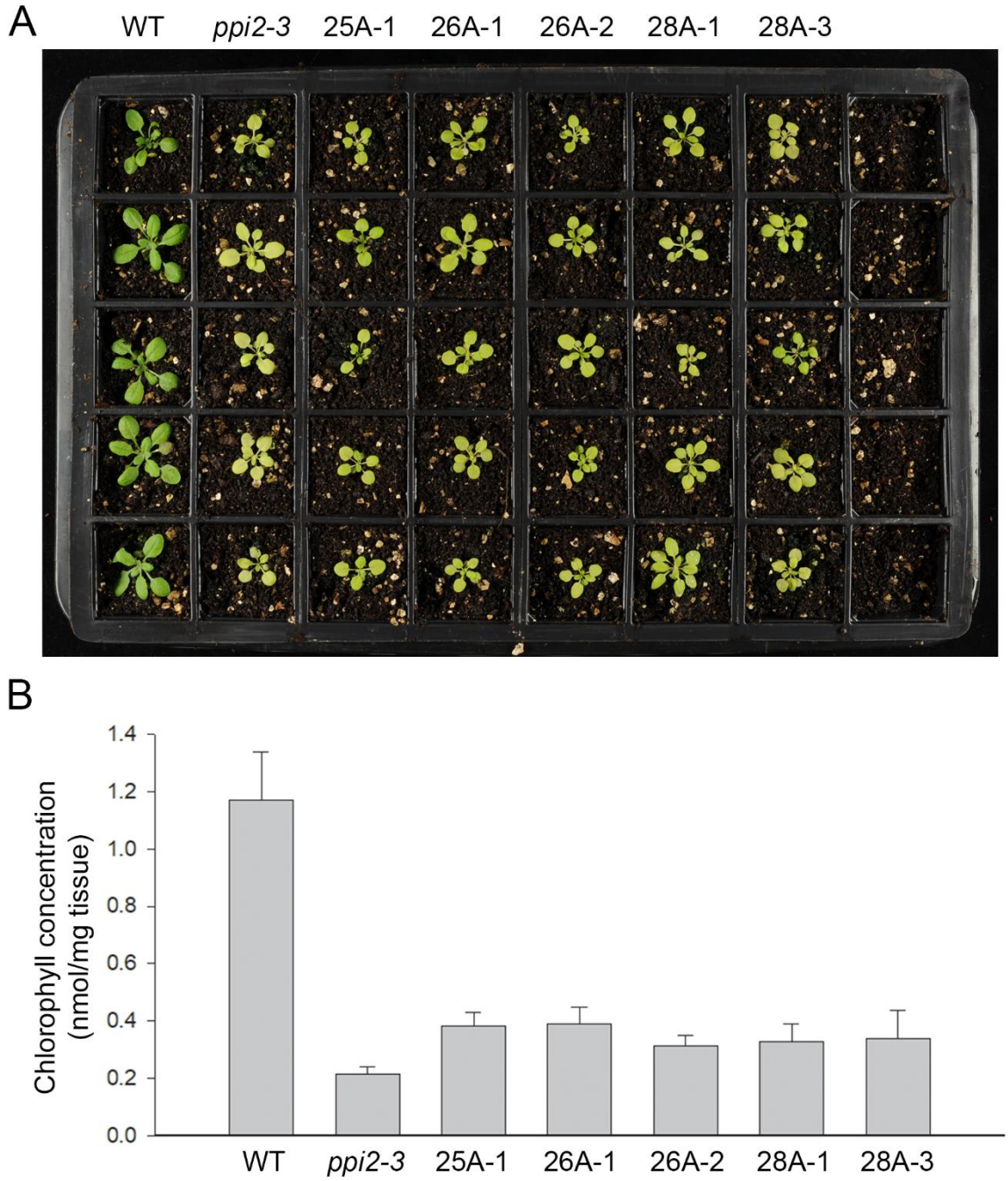


Figure 4.8

The sixth batch of new *ppi2-3* suppressor mutants.

(A) Visual appearance of the new *ppi2-3* suppressor mutants. The pictured plants were grown under identical conditions for 4 weeks.
(B) Quantification of chlorophyll accumulation in the new *ppi2-3* suppressor mutants. Measurements were taken on the day of photography (A) using a SPAD meter. Five plants were analysed per genotype. Error bars indicate standard deviation from the mean (n = 5).

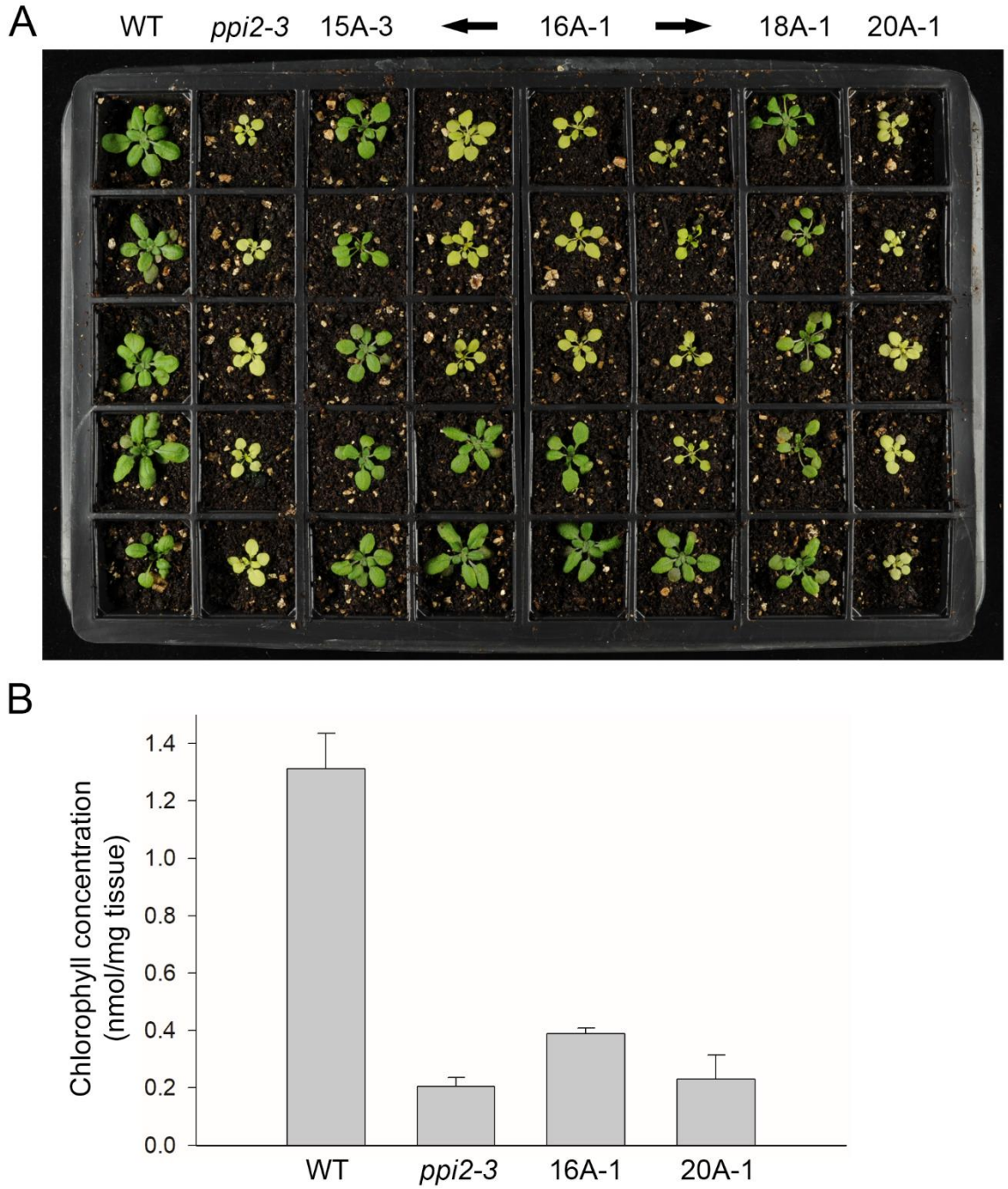


Figure 4.9

The seventh batch of new *ppi2-3* suppressor mutants – 'borderline' mutants.

(A) Visual appearance of the 'borderline' *ppi2-3* suppressor mutants. The pictured plants were grown under identical conditions for 4 weeks.
(B) Quantification of chlorophyll accumulation in selected *ppi2-3* suppressor mutants. Measurements were taken on the day of photography (A) using a SPAD meter. Five plants were analysed per genotype. Error bars indicate standard deviation from the mean (n = 3-4). Note: In the case of 16A-1, only plants clearly displaying the *ppi2-3* phenotype were analysed. In addition, 18A-1 plants were not analysed as they were presumed to accumulate wild-type levels of chlorophyll.

Table 4.1

A comprehensive catalogue of all of the new suppressor mutants described in this study, including information relating to the strength of suppression and the presence or absence of developmental defects.

Mutant name	**Batch number**	**Chlorophyll concentration (nmol/mg fresh tissue)**	**Variation (standard deviation)**	**Suppression strength (percentage increase in chlorophyll concentration relative to *ppi2-3*)**	**Phenotype comments**
1B-1	1	0.253	±0.068	+57.24%	Possibly early flowering (mild effect)
2A-2	1	0.325	±0.093	+102.39%	Leaf lesions in some plants
2A-3	1	0.308	±0.042	+91.81%	Dwarfed, small petioles
3A-1	1	0.430	±0.091	+167.56%	Early flowering
3A-2	1	0.260	±0.034	+61.5%	-
3A-4	1	0.289	±0.058	+79.99%	-
3A-5	2	0.446	±0.044	+154.40%	-
3B-1	2	0.240	±0.046	+37.25%	-
3B-2	2	0.265	±0.052	+51.25%	-
3B-3	2	0.288	±0.052	+64.19%	-
4A-1	2	0.285	±0.047	+62.44%	-

4A-2	2	0.406	±0.039	+131.95%	-
4B-1	3	0.313	±0.004	+63.05%	Possible dwarfism
4B-2	3	0.306	±0.038	+59.49%	Very mild leaf lesions
4B-4	3	0.310	±0.059	+61.79%	Possible dwarfism
6A-1	3	0.349	±0.093	+81.83%	Possible gigantism
6A-2	3	0.355	±0.028	+84.96%	Slightly variable phenotype
6A-3	3	0.282	±0.055	+47.26%	Severe dwarfism
7A-1	4	0.291	±0.016	+60.88%	Mild leaf edge defect
9A-1	4	0.309	±0.016	+70.61%	Suppression strength highest in young leaves
9A-2	4	0.410	±0.063	+126.63%	Leaf lesions in some plants
13A-1	4	0.345	±0.043	+90.91%	-
14A-1	4	0.307	±0.039	+69.85%	-
14A-2	4	0.741	±0.164	+309.73%	Possible mild dwarfism
15A-1	5	0.417	±0.048	+157.14%	-
15A-2	5	0.688	±0.083	+324.55%	Possible mild leaf edge defect
19A-1	5	0.381	±0.109	+135.15%	Mild leaf edge defect
21A-1	5	0.281	±0.019	+73.07%	Dwarfed
22A-1	5	0.378	±0.037	+133.44%	-

23A-1	5	0.292	±0.053	+80.35%	-
25A-1	6	0.383	±0.043	+78.14%	-
26A-1	6	0.388	±0.053	+80.84%	-
26A-2	6	0.312	±0.035	+45.10%	-
28A-1	6	0.328	±0.057	+52.53%	-
28A-3	6	0.339	±0.088	+57.70%	-
15A-3	7	-	-	-	Indistinguishable from wild type
16A-1	7	0.388	±0.017	+89.11%	Possibly contaminated with wild type seed
18A-1	7	0.327	±0.004	+59.24%	-
20A-1	7	0.231	±0.073	+12.32%	Very weak suppression

Table 4.2

A catalogue of the most promising suppressor mutants that will be taken forward in future work.

Mutant name	Revised name	Suppression strength	Notes
1B-1	*sop1*	+57.24%	-
3A-2	*sop2*	+61.5%	-
3A-4	*sop3*	+79.99%	May be the same genotype as 3A-2*
3A-5	*sop4*	+154.40%	-
3B-1	*sop5*	+37.25%	-
3B-2	*sop6*	+51.25%	-
3B-3	*sop7*	+64.19%	May be the same genotype as 3B-1*
4A-1	*sop8*	+62.44%	-
4A-2	*sop9*	+131.95%	-
4B-2	*sop10*	+59.49%	-
6A-2	*sop11*	+84.96%	-
7A-1	*sop12*	+60.88%	-
9A-1	*sop13*	+70.61%	-
13A-1	*sop14*	+90.91%	-
14A-1	*sop15*	+69.85%	-
14A-2	*sop16*	+309.73%	-
15A-1	*sop17*	+157.14%	-
15A-2	*sop18*	+324.55%	-
16A-1	*sop19*	+89.11%	Possibly contaminated with wild-type seed
18A-1	*sop20*	+59.24%	-
19A-1	*sop21*	+135.15%	-
22A-1	*sop22*	+133.44%	-
23A-1	*sop23*	+80.35%	-

25A-1	*sop24*	+78.14%	-
26A-1	*sop25*	+80.84%	-
26A-2	*sop26*	+45.10%	-
28A-1	*sop27*	+52.53%	-
28A-3	*sop28*	+57.70%	May be the same genotype as 28A-1*

** In these cases, two mutants with similar phenotypes were identified from the same batch of pooled seed. Thus, the two mutants may share a genotype, or they may represent two independent mutations.*

5 General discussion

5.1 Summary of key findings

This study can be divided into two general sections. In the first, the relationship between the SUMO system, plastid protein import, and the CHLORAD pathway was investigated. This involved a series of genetic, molecular, and biochemical experiments. In the second, a forward genetic screen for new regulators of plastid protein import was conducted. A total of 28 new mutants were identified and phenotypically characterised.

The relationship between SUMO, plastid protein import, and the CHLORAD pathway was first explored genetically. A number of SUMO system mutants were crossed with *ppi1*, a pale Toc33 null mutant, and the phenotypic consequences were examined. Interestingly, it was found that mutants representing various components of the SUMO system could partially suppress *ppi1* with respect to chlorosis (the plants were greener), and, in some cases, TOC protein accumulation and chloroplast development (the plants accumulated more TOC proteins and had better-developed chloroplasts). These mutants were: The E2 SUMO conjugating enzyme mutant *sce1-4*, the E3 SUMO ligase mutant *siz1-4*, and the SUMO protein mutants *sum1-1* and *sum2-1*. The minor SUMO isoform SUMO3 was also investigated, but the null mutant *sum3-1* was not found to suppress *ppi1*.

The suppression effects observed in some of the double mutants were demonstrated to be genetic background-specific phenotypes. The *sce1-4* or *siz1-4* single mutants were phenotypically characterised, and neither mutant displayed any increase in chlorophyll accumulation relative to a wild-type control. In addition, *sce1-4* was crossed with *tic40-4* and *hsp93-V-1*, two pale mutants that are associated with the TIC complex and, like *ppi1*, have defects in chloroplast protein import. The resulting double mutants were phenotypically characterised, and, remarkably, neither mutant displayed any change in chlorophyll accumulation relative to single mutant controls. These data supported the

notion that the SUMO mutations suppress lesions affecting the TOC complex specifically, in a manner similar to the well-characterised *sp1* and *sp2* mutations (Ling et al., 2012; Ling et al., 2019).

To complement the reverse genetic experiments, transgenic plants overexpressing SUMO1 or SUMO3 in the *ppi1* background were generated and phenotypically characterised. Interestingly, both sets of transgenic plants displayed an increase in the severity of the *ppi1* phenotype (the plants were paler and more severely stunted). While the effects observed in the SUMO1 overexpression plants were subtle, a large effect was observed in the SUMO3 overexpression plants. These data were consistent with a model in which the stability of the TOC complex is regulated by the SUMO system, with the principal involvement of SUMO3, and the minor involvement of SUMO1 and SUMO2.

The genetic experiments implied a functional link between the SUMO system and plastid protein import. This link was further explored in a series of molecular and biochemical experiments. The E2 SUMO conjugating enzyme SCE1 was found to physically interact with Toc159, Toc132, Toc34 and Toc33 in BiFC experiments, and with Toc159 and Toc132 in immunoprecipitation experiments. Similarly, the E3 SUMO ligases HPY2 and SIZ1 were found to physically interact with Toc159, Toc132, Toc34 and Toc33 in BiFC experiments, and HPY2 was found to interact with Toc159 in an immunoprecipitation experiment.

The protein-protein interaction studies demonstrated that key components of the SUMO system physically interact with the TOC complex. In an effort to demonstrate that TOC proteins are directly SUMOylated, two additional immunoprecipitation experiments were carried out. All three *Arabidopsis* SUMO isoforms were found to physically interact with Toc159 in a manner strongly suggestive of SUMOylation. SUMO3 most strongly interacted with Toc159, followed by SUMO2, and then SUMO1. SUMO3 alone was found to physically interact with Toc33 in a second experiment, and the immunoprecipitated

Toc33 appeared as multiple size-shifted bands on the membrane. This pattern was strongly indicative of SUMOylation.

In the second part of this project, a forward genetic screen was carried with the aim of identifying new regulators of plastid protein import. First, a recently identified Toc159 hypomorphic mutant, *ppi2-3*, was phenotypically characterised (Woodson et al., 2015; Ling et al., 2019). The *ppi2-3* mutation is located within the coding sequence of Toc159, and it reduces the stability of the encoded protein. The mutant is chlorotic but fertile, and the degree of chlorosis does not significantly vary between individuals. The *ppi2-3* mutant was therefore deemed to be a suitable starting genotype for a forward genetic screen. Around 29,000 *ppi2-3* seeds were chemically mutagenized and suppressor mutants (plants that appeared greener than their neighbours and non-mutagenised plants) were identified. After the exclusion of a number of mutants (on the grounds of infertility or growth defects), a total of 28 promising mutants were identified. All of these showed a robust increase in chlorophyll accumulation relative to the *ppi2-3* single mutant, and in all cases this suppression effect was stably transmitted to the M3 generation.

5.2 The relationship between SUMO, plastid protein import, and CHLORAD

The experiments described in this project demonstrated a genetic and functional link between plastid protein import and the SUMO system. In addition, the results of the genetic experiments were consistent with a link between the CHLORAD pathway and the SUMO system. The *ppi1*/SUMO system mutants closely resembled the *ppi1 sp1* and *ppi1 sp2* double mutants (Ling et al., 2012; Ling et al., 2019), which implies that the SUMO system may function in the CHLORAD pathway; however, this has not yet been directly demonstrated.

Although the genetic and molecular experiments demonstrated a general functional link between the SUMO system and plastid protein import (and, possibly, the CHLORAD pathway), the precise, mechanistic details of this link were not elucidated. A number of possible models have been considered, as follows.

5.2.1 Does SUMO recruit Cdc48 to the TOC complex?

As discussed in Chapter 1, Cdc48 associates with two major cofactors, Ufd1 and Npl4 (Baek et al., 2013). These cofactors recruit Cdc48 to specific substrates. Interestingly, in baker's yeast, Ufd1 contains a SUMO-interaction motif (SIM), and has been shown to non-covalently bind SUMO moieties (Nie et al., 2012). Moreover, the SUMO-mediated recruitment of Cdc48 has important roles within the maintenance of genome stability in yeast (Bergink et al., 2013).

The results presented in this study are consistent with a model in which TOC proteins are first ubiquitinated by SP1, and then SUMOylated (in response to stress and/or developmental triggers); and in which such a dual SUMOylation/ubiquitination process promotes the recruitment of Cdc48 via the Ufd1 cofactor. In this model, Cdc48 would then act to remove the TOC proteins from the chloroplast outer membrane in concert with SP2 (Ling et al., 2019), so that the proteins may then degraded in the cytosol via the 26S proteasome.

To test this model, *ufd1* mutants could be obtained and crossed with *ppi1*. A phenotypic suppression effect would indicate the involvement of Ufd1 in the regulation of plastid protein import. Next, the double mutant could be transgenically complemented with two constructs: The first would encode the wild-type Ufd1 protein, and the second would encode a Ufd1 variant bearing mutations in its SIM domain. If Ufd1 does guide Cdc48 to the SUMOylated TOC complex, the *ppi1 ufd1* double mutant that was complemented with the wild-type Ufd1 would be expected to show a reversion to the *ppi1* phenotype, whereas such reversion would be blocked in the *ppi1 ufd1* double mutant complemented with the SIM-mutated Ufd1 variant.

Alongside the genetic experiments, some immunoprecipitation experiments would also be informative. First, the physical interaction between Ufd1 and TOC proteins could be explored. The ability of Ufd1, and also a SIM-mutated Ufd1 variant, to pull down TOC proteins could be determined. In addition, the physical interaction between Cdc48 and TOC proteins could be explored in protoplast-based immunoprecipitation experiments. The protoplasts could be isolated from a range of genetic backgrounds (for example, *ufd1*, *sce1-4*, SUMO1 OX, SUMO3 OX etc), in order to gauge how different factors and processes influence the interaction. The physical interaction between Cdc48 and TOC proteins was demonstrated in previous work (Ling et al., 2019). The genetic perturbation of SUMOylation would be expected to modify the ability of Cdc48 to efficiently immunoprecipitate TOC proteins. For example, given that my data point to a significant role for SUMO3 in TOC regulation, in the SUMO3 OX protoplasts, Cdc48 would be expected to pull down TOC proteins with greater efficiency than in wild-type protoplasts; and the inverse would be expected in protoplasts isolated from SUMO mutants.

5.2.2 Does SP1 act as a STUbL?

STUbLs are a class of E3 ubiquitin ligases that specifically target polySUMOylated proteins (Sriramachandran and Dohmen, 2014). As a result of their activity, the SUMO chains themselves are ubiquitinated. As SP1 ubiquitinates proteins that are also SUMOylated, it is possible that SP1 is a STUbL. If this is the case, the results described in this project could be explained via a model in which TOC proteins are SUMOylated, and this SUMOylation is recognised by SP1. SP1 would then ubiquitinate the TOC proteins, and SP2 and Cdc48 would extract the modified proteins from the chloroplast outer membrane as in the published CHLORAD scheme (Ling et al., 2019). It is important to note that this model is compatible with the first model described above (see Section 5.2.1); i.e. SUMO could promote both SP1 activity and Cdc48 recruitment.

SP1 is a small protein with a simple domain architecture. All functionally characterised STUbLs contain two key domains: A RING domain, and one or more SIM

domains, which are essential for the recognition of SUMOylated proteins. SP1 has a RING domain, but no obvious SIM domains. Most STUbLs contain multiple SIM domains which enable them to recognise SUMO chains (Sriramachandran and Dohmen, 2014). SP1's lack of obvious SIM domains may make it an unlikely STUbL candidate. However, it could be that SP1 physically associates with cofactors that enable it to recognise SUMOylated proteins, though there is no literature precedent for this – all characterised STUbLs contain SIM domains (Sriramachandran and Dohmen, 2014).

To test this model, a number of simple experiments could be carried out. First, the E4 SUMO ligase mutants *pial1* and *pial2* could be crossed with *ppi1*, and the resulting double and triple mutants phenotypically characterised. Suppression effects would indicate that polySUMOylation (a prerequisite for STUbL action) is important in the regulation of plastid protein import. In addition, the physical interaction between SP1 and TOC proteins could be explored in protoplast-based immunoprecipitation experiments, using protoplasts isolated from different genetic backgrounds including SUMO overexpression lines and mutants. It is well established that SP1 associates with TOC proteins (Ling et al., 2012); however, it is not known if these interactions are SUMO dependent. In the protoplast assay, if the strength of each SP1-TOC interaction was found to be modified in SUMO mutant and overexpression backgrounds, this would indicate a role for SUMO.

5.2.3 Is SP1 SUMOylated?

As already noted, the *ppi1*/SUMO double mutants showed a remarkable similarity with the *ppi1 sp1* and *ppi1 sp2* double mutants (Ling et al., 2012; Ling et al., 2019). Thus, perturbing SUMOylation or the CHLORAD pathway suppresses *ppi1* with respect to chlorophyll accumulation, TOC protein abundance, and chloroplast development. The effects were smaller than the effects mediated by *sp1* and *sp2* null mutants, but this can be explained by the fact that none of the SUMO system mutants used in this analysis are totally deficient in SUMOylation. The simplest, but most speculative, explanation for this

similarity is that CHLORAD components are, themselves SUMOylated. SP1 is, in unstressed plants, present at extremely low abundance within the cell. This is a consequence of SP1's ability to autoubiquitinate, a behaviour commonly observed in RING-type E3 ubiquitin ligases (Deshaies and Joazeiro, 2009; Ling et al., 2012). Under conditions in which the CHLORAD pathway is activated, SP1 must be able increase in abundance and/or activity. This could be achieved through the SUMOylation of SP1's exposed lysine residues. The SUMOylation of these lysine residues could prevent SP1's ubiquitination, allowing it to rapidly increase in abundance and promote the turnover of TOC proteins.

This model does not explain the significance of the SUMOylation of TOC proteins, but it is certainly compatible with the other two models presented in this section. Although this model is speculative, some supporting evidence has emerged in the last year. For example, in a biochemical screen for proteins that physically associate with SP1, a SUMO protease was identified (Yi Sun and Paul Jarvis, unpublished data).

Some simple experiments could be carried out to test this model. Protein-protein interaction studies could determine whether SP1 physically interacts with the SUMO system, and whether SP1 is SUMOylated. BiFC and immunoprecipitation experiments similar to the ones described in this project could be carried out. In addition, the abundance of SP1 could be tested in various genetic backgrounds (i.e. SUMO overexpression lines, SUMO mutants). If SP1 is protected from turnover by SUMO, SP1 would be expected to increase in abundance when the rate of SUMOylation was increased.

5.3 New suppressors of Toc159 mutant *ppi2-3*

The forward genetic screen for new regulators of plastid protein import yielded 28 novel mutants, which I have named *suppressor of ppi2-3 1* to *28* (*sop1-28*). As none of the causative loci have been identified at the time of writing, it is not yet clear where this

project will lead. Genes affecting the stability, activity, localisation or function of TOC proteins may be identified. In addition, new components and regulators of the CHLORAD pathway may be uncovered. The recent discovery of the CHLORAD pathway demonstrates the power of this general approach (Ling et al., 2012; Ling et al., 2019). The future direction of the project is outlined in detail in Chapter 4.

5.4 Concluding remarks

The work presented in this project has uncovered a previously unknown genetic and functional link between the SUMO system and plastid protein import. In addition, these findings are of broad significance as they represent the first example of the SUMOylation of proteins at the chloroplast outer membrane. This work therefore represents an important general advance in the chloroplast field, as well as a narrower advance in our understanding of the regulation of plastid protein import, and, perhaps, the CHLORAD pathway. The SUMO work is complemented by the forward genetic screen, which should also advance our understanding of the regulation of plastid protein import and, possibly, the CHLORAD pathway.

6 Materials and methods

6.1 Plant growth conditions and physiological studies

The *Arabidopsis thaliana* Columbia-0 (Col-0) ecotype was used throughout the study. In most cases, plants were grown in a growth chamber (Percival Scientific, Inc.) under long-day conditions (16 hours light, 8 hours dark). The light intensity was approximately 120 µE m^{-2} s^{-1}, and the temperature was held at a constant 20°C with a humidity of ~70%. In a few cases, plants were grown in a light-supplemented glasshouse under standard long-day conditions.

6.1.1 Plants grown on media

Plants were grown on Murashige and Skoog (MS) media supplemented with 0.5% (w/v) sucrose and 0.8% (w/v) phyto-agar. The MS was made up as follows: 4.3 g/L basal salt mixture (Duchefa Biochemie), 0.5% (w/v) sucrose, 0.05% (w/v) 2-(N-morpholino)ethanesulfonic acid (MES). After the addition of phyto-agar, the media was adjusted to pH 5.7 via the addition of potassium hydroxide (KOH) and autoclaved at 115°C for 15 minutes. The liquid MS media was allowed to cool to ~60°C and then poured into standard 9 cm diameter petri dishes to a depth of 1 cm under sterile conditions in a laminar flow hood and allowed to set for 20-30 minutes with the lids open.

Prior to sowing, seeds were surface-sterilised via a 5-minute incubation in ethanol-based sterilisation solution (70% (v/v) ethanol, 0.05 % (v/v) Triton X-100) with gentle shaking, followed by a 10-minute incubation in absolute ethanol, also with gentle shaking. After sterilisation, seeds were transferred to ethanol-soaked filter paper and left to dry in a laminar flow hood for ~20 minutes. After drying, seeds were sown directly onto plates and stratified at 4°C for 2-3 days in the dark. In most cases, plants were used in downstream experiments after 10-14 days of growth.

6.1.2 Plants grown on soil

When growth to later stages or maturity was required, plants were grown on compost containing 80% (w/w) soil (Levington Pot and Bedding M2), 20% vermiculite (Sinclair, medium particle size), and 0.3 g/L Imidasect (Sinclair, an insecticide). For most of the analyses, plants were grown in module trays containing 24 or 40 cells and watered as required from below. In most cases, seeds were sown directly onto wetted soil and stratified at 4°C for 2-4 days before transfer to the growth chamber or greenhouse. However, in some cases, seedlings were transferred from MS plates to soil after 2-3 weeks of growth.

6.1.3 Crossing plants

Plants were grown on soil and allowed to bolt. In most cases, crosses were made between plastid import mutants with pale, recessive phenotypes and phenotypically normal SUMO mutants. The plastid import mutants were used as female ‘acceptor’ plants so that the absence of the pale recessive phenotype in the F1 generation could be used as an indication that the cross had been successful. ‘Acceptor’ flowers were emasculated under a magnifying glass using fine-tipped forceps, and then 3-4 ‘male donor’ flowers were removed and used to fertilise the exposed pistils of the ‘acceptor’ flowers. The siliques arising from successful crosses were harvested after ~10 days of development.

6.1.4 Chlorophyll measurements

Chlorophyll measurements were taken from mature rosette leaves in each instance. Plants were typically allowed to grow for 3-5 weeks before measurements were taken. A Konica-Minolta SPAD-502 meter was used to take each measurement. The SPAD meter generates raw spectrophotometric readings; these were converted into chlorophyll concentrations (nmol/mg fresh weight) via published conversion equations (Ling et al., 2011).

6.2 Plant DNA and RNA work

6.2.1 DNA extraction from plant tissue

All centrifugation steps were performed using a microcentrifuge (Eppendorf centrifuge, 5415 D, rotor EL 082). DNA was isolated from single, mature rosette leaves. Each leaf was flash-frozen in liquid nitrogen and ground into a fine powder by hand using a pestle and mortar or via the use of a TissueLyser (Qiagen). Where the TissueLyser was used, each sample was homogenised at 20-24 Hz for 1-2 minutes in a 2 mL microcentrifuge tube containing an 5 mm diameter steel ball. An aliquot (0.5 mL) of DNA extraction buffer (200 mM Tris(hydroxymethyl)aminomethane-HCl (Tris-HCl) pH 7.5, 250 mM NaCl, 25 mM (ethylenedinitrilo)tetraacetic acid (EDTA), 0.5% (w/v) sodium dodecyl sulfate (SDS)) was added to each sample. Each sample was then vortexed for 10 seconds and centrifuged for 5 minutes (14,000 × *g*, 4°C). The supernatant of each sample was then transferred to a fresh tube containing 450 µL isopropanol and mixed thoroughly via inversion. The samples were stored at -20°C for between 30 minutes and 2 hours and centrifuged for 10 minutes (14,000 × *g*, 4°C). The supernatant was then removed from each sample and the pellets were washed with 1 mL 70% (v/v) ice-cold ethanol. The samples were centrifuged for 5 minutes (14,000 × *g*, 4°C). The ethanol was then removed from each tube, and each pellet was air-dried for ~30 minutes at room temperature. Each pellet was resuspended in 50µL dH_2O and stored at 4°C overnight. The next day, the samples were centrifuged (15,000 × *g*, 4°C) and the pellets were resuspended in 25-50µL TE buffer (10 mM Tris-HCL pH 8.0, 1 mM EDTA pH 8.0).

6.2.2 PCR and PCR genotyping

Polymerase chain reaction (PCR) was carried out via the use of a T3 Biometra Thermocycler. In most cases, standard reaction solutions were used (10 mM Tris-HCL pH 8.3, 50 mM KCl, 50 mM MgCl, 0.05% (v/v) Tween 20, 200 µM of each deoxyribonucleoside triphosphate (dNTP), 125 nM Primer 1, 125nM Primer 2, 0.5 units of Taq DNA Polymerase). The Biometra programme used varied according to primer

properties, expected amplicon size, and application. For example, the annealing step temperature was increased for primer pairs with high melting temperatures and the extension step duration was increased for large amplicons. In the case of cloning PCR, Phusion proof-reading polymerase (NEB) was used to minimise amplification errors. In these cases, reactions were set up in accordance with the manufacturer's instructions.

A series of *A. thaliana* T-DNA insertion mutants were used in this study. A catalogue of these mutants can be found at the beginning of Chapter 2 (Table 2.1). Each mutant was genotyped via PCR. Primers specific to each gene of interest were generated, and the primers were carefully positioned within the genomic DNA (gDNA) template such that one of the primers could generate a product of approximately 500 bp when paired with standard left border genotyping primers that anneal to the T-DNA insert (e.g. LBb1.3 for SALK lines). In each genotyping procedure, two parallel PCR reactions were set up using DNA extracted from the mutant as a template. In the first, the two *A. thaliana* 'gene specific' primers were used, and in the second, one 'gene specific' primer was paired with the appropriate T-DNA left border primer (see Table 6.1). The two products were resolved via agarose gel electrophoresis, and the absence of a 'gene specific' band alongside the presence of a 'left border' band was taken to indicate the homozygous presence of a T-DNA insertion.

6.2.3 Agarose gel electrophoresis

DNA samples were run on 1% (w/v) agarose gels. Each gel was cast by adding 1% (w/v) agarose to an appropriate volume of 0.5× TBE buffer (40 mM Tris-HCl pH 8.3, 45 mM boric acid, 1 mM EDTA). The solution was then microwaved to solubilise the suspended agarose. After the solution had cooled to ~60°C, an appropriate volume (1 µL/50 mL) of 1× SYBR Safe DNA stain (Invitrogen) was added. The solution was then transferred to a gel casting tray and allowed to set. DNA samples were loaded in 1× Orange G loading buffer (0.1% (w/v) Orange G, 50% (v/v) glycerol) and run in a tank (Mini-Sub Cell GT, Bio-Rad) containing 0.5× TBE buffer at ~10 V/cm for 20-30 minutes,

or until the Orange G dye had migrated beyond the mid-point of the gel. A DNA ladder was added to the first well of each gel (1kb+, ThermoFisher). Each gel was visualised via use of a GeneFlash (Syngene) or LAS-4000 (GE Healthcare) bioimager.

6.2.4 RNA extraction from plant tissue

Single rosette leaves were taken from 3-4-week-old plants and transferred to 2mL Eppendorf tubes each containing a 5 mm diameter steel bead. Each sample was flash-frozen and homogenized via use of a TissueLyser (Qiagen) under standard conditions (see Section 6.2.1). Total cellular RNA was extracted from the homogenate via use of a Spectrum Plant Total RNA Kit (Sigma) in accordance with the manufacturer's instructions. Contaminating DNA was removed from each sample during the procedure via the inclusion of an optional, 'on column' DNase treatment step (On-Column DNase I Digestion Set, Sigma). The concentration of each RNA sample was determined via spectrophotometric analysis using a NanoPhotometer P330 (Implen).

6.2.5 First-strand cDNA synthesis

Plant total RNA samples were used as a template for complementary DNA (cDNA) synthesis. In all cases, the Superscript IV system (Invitrogen) was used in accordance with the manufacturer's instructions. An equal amount of each RNA sample (2-3 µg) was used per 20 µL reaction, and then each cDNA sample was diluted 1:5 or 1:10 before use in reverse transcription polymerase chain reaction (RT-PCR) applications.

6.2.6 Semi-quantitative RT-PCR

Total RNA samples were taken from representative plants of each genotype. Each sample was used as a template for first-strand cDNA synthesis as described in the previous section; however, prior to synthesis, the amount of template RNA included in each reaction was normalised with respect to the sample with the lowest concentration (ng/µL). Each cDNA sample was then used as a template for two PCR reactions, one using primers specific to the gene of interest, and one using primers specific to a control

gene *EUKARYOTIC TRANSLATION INITIATION FACTOR 4E-1* (*eIF4E1*). The eIF4E1 gene (AT4G18040) was used as a control as it shows uniform expression throughout the plant (Rodriguez et al., 1998). All primers used in these analyses are detailed in Table 6.4. Each PCR reaction was carried out under standard conditions but with a relatively small number of cycles (25-26). The small number of cycles ensured that the samples were captured during the exponential phase of amplification, making the procedure semi-quantitative. Each sample was resolved on an agarose gel and imaged following staining using the LAS-4000 bioimager (see Section 6.2.3).

6.3 Bacterial work and molecular cloning

6.3.1 *E. coli* strains, growth conditions, and DNA isolation

Escherichia coli was grown in solution in Lysogeny broth (LB; 10% (w/v) tryptone, 5% (w/v) yeast extract, 5% (w/v) NaCl) (Sambrook, 2001). The LB was autoclaved at 120°C before use. *E. coli* was also grown on LB agar plates; to make these, 1% (w/v) agar was added to liquid LB and the solution was autoclaved similarly at 120°C, poured into 9 cm diameter petri dishes to a depth of 1 cm, and allowed to set. When grown in solution, cultures were incubated at 37°C with vigorous shaking (200 rpm). When streaked on plates, cultures were incubated at 37°C without shaking.

The same variety of competent cells was used in all cloning procedures (α-Select Silver Efficiency Chemically Competent Cells, Bioline). The cells were transformed via heat shock in a modified version of the manufacturer's protocol. First, 10 µL aliquots of competent cells were defrosted on ice. Next, 1 µL plasmid DNA was added to each tube and the contents of the tubes were mixed via gentle flicking. The samples were incubated on ice for 30 minutes, then heat shocked at 42°C for exactly 30 seconds. The tubes were then incubated on ice for 2 minutes, and then 490 µL LB was added to each sample. The tubes were incubated at 37°C for 1 hour with vigorous shaking (200 rpm). After this step, cells were subjected to antibiotic selection via growth on LB agar plates containing the relevant antibiotic at the following concentrations: 100 µg/mL ampicillin, 50 µg/mL

kanamycin, 50 µg/mL spectinomycin, and 10 µg/mL gentamycin. The cells were embedded within the plates via use of disposable plastic spreaders.

Plasmid DNA was isolated from selected bacterial strains via the use of a miniprep kit (Qiagen Plasmid Mini Kit) or a midiprep kit (Qiagen Plasmid Midi Kit). For miniprep extractions, 5-10 mL overnight cultures were used, and 50-200 mL overnight cultures were used in midiprep extractions. In both cases, the manufacturer's protocol was followed. Midiprep extractions were performed when high concentrations of plasmid DNA (>1µg/µL) were required. DNA concentrations were measured via spectrophotometric analysis using a NanoPhotometer P330 (Implen).

6.3.2 *Agrobacterium* growth conditions and transformation

All transgenic plants used in this study were transformed via the *Agrobacterium tumefaciens*-based 'floral dip' method (Clough and Bent, 1998). The GV3101 strain of *Agrobacterium* was used in all cases. A microcentrifuge was used in all centrifugation steps (Eppendorf centrifuge, 5415 D, rotor EL 082). The *SUMO1* and *SUMO3* coding sequences were cloned into the pH2GW7 vector (Karimi et al., 2002) via the Gateway system (described in detail in Section 6.3.4 below). The finalised constructs were used to transform *Agrobacterium* cultures via the freeze-thaw method (Hofgen and Willmitzer, 1988). In each transformation, 1 µg of plasmid DNA was added to a 1.5 mL microcentrifuge tube containing a 100 µL aliquot of competent cells. The tubes were mixed and flash-frozen in liquid nitrogen. The tubes were then removed and allowed to defrost at 37°C (5 minutes). To each culture, 1 mL of LB medium was then added, and the tubes were incubated at 28°C (4 hours with shaking at 200 rpm). The transformed cells were then collected via centrifugation (1 minute, 7000 × *g*), and resuspended in approximately 100 µL of residual LB. The cells were plated onto selective medium (spectinomycin, 50 µg/mL; gentamycin, 10 µg/mL) and incubated at 28°C for two days. Individual colonies were then picked using a sterile pipette tips and used to inoculate aliquots of liquid LB medium containing the appropriate antibiotics (spectinomycin, 50

μg/mL; gentamycin, 10 μg/mL); cultures were grown overnight at 28°C with shaking (220 rpm). The next day, the cells were collected via centrifugation (10 minutes, 6000 × *g*) and resuspended in a 5% (w/v) sucrose solution. At this point, the solution was at an approximate optical density of 0.8 (OD_{600}).

Plants were grown on soil for approximately five weeks prior to transformation. Once the plants had grown to an approximate height of 5 cm, the primary inflorescences were removed to encourage the growth of many secondary inflorescences. On the day of the transformation, all mature flowers and siliques were removed using fine tweezers.

The sucrose solutions containing the transformed *Agrobacterium* cultures were supplemented with 300 μL/L Silwet L-77 (Lehle Seeds), and all aerial parts of the plants to be transformed were inverted and dipped in the solutions for approximately 30 seconds. The dipped plants were transferred to the greenhouse and grown under plastid propagator lids for one day. The lids were then removed, and the plants were allowed to flower and set seed. The T1 seed was collected and sown onto MS media supplemented with appropriate antibiotics (spectinomycin, 50 μg/mL; gentamycin, 10 μg/mL). Surviving plants were allowed to set seed, and the resulting T2 seeds were analysed in another round of antibiotic selection. Single insert lines were identified on the basis of the segregation ratio of resistant/sensitive plants (a 3:1 ratio was sought). Single insert plants were allowed to flower and set seed, and homozygous lines were identified in the T3 generation, again on the basis of the segregation of resistance marker (the progeny of homozygous plants showed 100% antibiotic resistance).

6.3.3 Restriction enzyme-based cloning

The constructs used in the bimolecular fluorescence complementation (BiFC) experiments were generated via restriction cloning. The *SCE1*, *SIZ1* and *HPY2* coding sequences were cloned into the vector pSAT4A-nEYFP-N1 (Tzfira et al., 2005). The TOC and negative control constructs were generated by a previous member of the group (Dr Feijie Wu). In the case of ΔOEP7, the first 105 base pairs of the *OEP7* coding sequence

were amplified. This truncated sequence is sufficient to efficiently target the full-length YFP protein to the chloroplast outer envelope (Lee et al., 2001). The primers used in each cloning project are described in Table 6.2, and these were designed as follows. The coding sequence of each gene to be cloned was assessed for the presence of restriction sites, and pairs of restriction enzymes that were a) unable to cut each coding sequence and b) had restriction sites represented in the pSAT4A-nEYFP-N1 vector's multiple cloning site were selected. Primer sequences specific to each coding sequence were selected (these were typically 20 nucleotides in length, and were chosen to anneal at the start and stop codons), and to these were appended 5' extensions containing recognition sequences for the restriction enzymes to be used during cloning into the pSAT4A-nEYFP-N1 vector. Each coding sequence was amplified via Phusion PCR. The amplification of each PCR product was checked via agarose gel electrophoresis and, where amplification was successful, the bands were cut out and purified via use of a GenElute Gel Extraction Kit (Sigma). The purified products were then digested with the appropriate restriction enzyme pair in accordance with the manufacturer's instructions (New England Biolabs). The pSAT4A-nEYFP-N1 vector was digested in parallel. The digested fragments and vectors were then separated via agarose gel electrophoresis, gel-purified via use of a GenElute Gel Extraction Kit (Sigma) and ligated via incubation with T4 DNA ligase (NEB) in accordance with the manufacturer's instructions. Each reaction was used to transform competent *E. coli* cells (as outlined in Section 6.3.1) and the transformed cells were streaked onto plates containing the appropriate antibiotic (ampicillin, 100 µg/mL). The next day, single colonies were selected and each one was used to inoculate a 1-5 mL volume of LB containing the appropriate antibiotic in 15 mL culture tubes. Cultures were grown overnight and subjected to plasmid isolation via use of a miniprep kit (as described in Section 6.3.1). The plasmid DNA of each colony was sent for commercial Sanger sequencing (Source Bioscience), using the cloning primers as sequencing primers, and the sequence data was analysed using the Chromas and SnapGene software packages.

Constructs bearing accurate, in-frame insertions were retained for use in BiFC experiments.

6.3.4 Gateway cloning

The constructs used in the SCE1, HPY2 and SIZ1 immunoprecipitation (IP) experiments were generated via the Gateway cloning system (Invitrogen), as were the SUMO1, SUMO2, and SUMO3 constructs used in the SUMOylation assays, and the SUMO1 and SUMO3 constructs used in the plant transformation procedures. Primers specific to the cloning sequence of each gene were generated (these were typically 20 nucleotides in length and were chosen to anneal at the start and stop codons) and appended with 5' AttB1 (forward primer) and AttB2 (reverse primer) sequences (see Table 6.3). Each gene was amplified via Phusion PCR and inserted into either the pDONR201 entry vector (for SCE1, HPY2 and SIZ1 for use in immunoprecipitation experiments; and for SUMO1 and SUMO3 for use in the plant transformations) or the pDONR221 entry vector (for SUMO1, SUMO2 and SUMO3 for use in immunoprecipitation experiments), via incubation with BP clonase (reactions proceeded according to the manufacturer's instructions). The reaction solutions were then used to transform competent *E. coli* cells, and the transformed cells were plated on selective medium (Kanamycin, 50 µg/mL in both cases). The next day, single colonies were selected and each one was used to inoculate a 1-5 mL volume of LB containing an appropriate antibiotic in 15 mL culture tubes. Plasmid DNA was isolated from each culture and the DNA samples were sent for commercial Sanger sequencing (Source Bioscience) using the cloning primers as sequencing primers. The sequence data was analysed using the Chromas and SnapGene software packages, and constructs bearing accurate insertions were retained.

The DNA samples from selected entry clones were then combined with LR clonase and destination vector DNA: p2GWY7 (SCE1, HPY2, SIZ1), p2YGW7 (SUMO1, SUMO2, SUMO3) and pH2GW7 (SUMO1, SUMO3). The reactions were set up in accordance with

the manufacturer's instructions. The inserts were transferred from the entry clone to the destination vector via the action of LR clonase. Each reaction was then used to transform competent cells, and the cells were subjected to antibiotic selection. The next day, single colonies were selected and each one was used to inoculate a 1-5 mL volume of LB containing the appropriate antibiotic in 15 mL culture tubes. Plasmid DNA was isolated from each culture and the DNA samples were sent for commercial Sanger sequencing. Constructs bearing accurate insertions were retained for use in immunoprecipitation experiments.

6.4 Protein analysis

6.4.1 Protein extraction

Rosette leaves were taken from 3-4-week-old plants and flash-frozen in liquid nitrogen in 2 mL microcentrifuge tubes containing 5 mm diameter steel beads. Approximately 50-60 mg of tissue was taken in each case. Each sample was homogenised using a TissueLyser (20 Hz for 1-2 minutes) and suspended in 250-300 µL protein extraction buffer (PEB; 100 mM Tris-HCl pH 6.8, 10% (v/v) glycerol, 1% (w/v) SDS, 0.1% (v/v) Triton X-100, 5 mM EDTA). The PEB was supplemented with 10 mM dithiothreitol (DTT) and 10 µL/mL protease inhibitor cocktail (PIC; Roche) immediately prior to use. Each sample was thoroughly vortexed and incubated on ice for 20 minutes. The samples were then centrifuged for 20 minutes in a microcentrifuge (15,000 × *g*, 4°C) and each supernatant was transferred to a fresh tube. Where necessary, the protein concentration of each sample was determined via a Bradford assay. Before analysis, a standard curve was generated using a series of Bovine Serum Albumin (BSA; Sigma) standards (1, 2, 4, 8 and 10 µg/µL). Each 1mL standard reaction contained 1 µL of protein sample, 799 µL distilled H_2O (dH_2O), and 200 µL Protein Assay dye (Bio-Rad). Absorbance readings (OD_{595}) were taken via a NanoPhotometer P330 (Implen). The plant protein samples were then prepared and processed in the same way, and the spectrophotometric readings were converted into protein concentration values (µg/µL) by

reference to the standard curve using Microsoft Excel. Prior to analysis via sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE), each sample was diluted 1:1 in 2× Laemmli buffer (Laemmli, 1970) (0.1 M Tris-HCl pH 6.8, 2% (w/v) SDS, 20% (v/v) glycerol, 0.02% bromophenol blue, 0.1 M DTT) and boiled for 5 minutes at 100°C in screw-cap 1.5 mL microcentrifuge tubes.

6.4.2 SDS-PAGE

Protein samples in Laemmli buffer were loaded onto polyacrylamide gels and analysed via SDS-PAGE (Laemmli, 1970). The acrylamide concentration of each gel depended on the size of the proteins to be resolved. For important experiments (i.e. those that were expected to yield publication-quality data), precast gradient gels were often used (4–15% Mini-PROTEAN TGX gels, Bio-Rad). However, for most purposes, gels were manually cast using glass plates and a dedicated casting assembly (Mini Protean III, Bio-Rad). The resolving gel solution was prepared via the addition of a variable proportion (26-40% (v/v)) of ProtoGel (Geneflow) to 1× Resolving buffer (375 mM Tris-HCl pH 8.8, 0.1% (w/v) SDS, 0.5% (w/v) ammonium persulphate (APS), 0.05% (v/v) tetramethylethylenediamine (TEMED)). ProtoGel contains 30% (v/v) acrylamide and 0.8% (v/v) bis-acrylamide. The variable proportion of ProtoGel was used to create gels that contained 8-12% (v/v) acrylamide. The APS and TEMED were added last as they rapidly initiate polymerization. The resulting solution was transferred to prepared glass plates (separated by 0.75 mm spacers) and immediately overlaid with dH_2O to ensure efficient polymerization. After 30 minutes (at room temperature), the overlaid water was carefully removed and replaced with stacking solution (16.75% (v/v) Protogel, 1× Stacking buffer (0.125 M Tris-HCl pH 6.8, 0.1% (w/v) SDS, 0.75% (w/v) APS, 0.125% (v/v) TEMED). A well-forming comb was immediately inserted, and the gel was left to polymerize for 30 minutes at room temperature.

Each gel was inserted into an electrophoresis chamber (Mini-Protean III, Bio-Rad) and the inner chamber of the tank was filled with Tris-Glycine-SDS running buffer (TGS;

25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS) such that the wells of each gel were fully submerged. Running buffer was poured into the outer chamber of the tank until it was half-full. Each protein sample was centrifuged for 3-10 minutes in a microcentrifuge (16,000 × *g*, room temperature) and 2-20 µL of sample was loaded per well depending on the requirements of the experiment. A pre-stained protein ladder was loaded in the first well in each case (Precision Plus Protein Dual Color Standards, Bio-Rad). The gels were then run at 180 V at room temperature for 60-90 minutes in most cases, or, in a small number of experiments, for 4-5 hours.

6.4.3 Coomassie staining

In select cases, SDS-PAGE gels were incubated in InstantBlue Ultrafast Protein Stain (Sigma) for 10-20 minutes (room temperature with gentle shaking; ~ 150 rpm). The stained membranes were then photographed using a desktop scanner and two acetate sheets.

6.4.4 Western blotting

After SDS-PAGE, separated proteins were usually transferred to nitrocellulose blotting membranes (Amersham Protran 0.2 NC nitrocellulose Western blotting membrane, GE Healthcare). SDS-PAGE gels were overlaid with membrane and 'sandwiched' into a transfer cassette. Each cassette was loaded into a transfer tank (Mini Trans-Blot Module, Bio-Rad) filled with ice-cold transfer buffer (25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS, 20% (v/v) methanol). The transfer conditions were the same in each case: 100 V for 60 minutes with stirring. The transfer buffer was chilled to 4°C before use, and an ice pack was inserted into the transfer tank to keep the assembly cool.

After the transfer had finished, the membranes were removed from the cassettes and incubated in blocking solution for 30 minutes at room temperature with gentle shaking (~100 rpm). Blotto (5-8% (w/v) skimmed milk powder suspended in Tris-buffered saline-Tween solution (TBST-T; 10 mM Tris-HCl pH 8, 150 mM NaCl, 0.05% (v/v) Tween 20))

was used in each instance. After blocking, each membrane was cut into strips corresponding to protein size ranges of interest, and incubated with primary antibodies overnight at 4°C. Each primary antibody was diluted in 15-20 mL Blotto, and the dilution ratio depended on the properties of the individual antibodies. The next day, the membrane strips were washed three times, each time with 20 mL TBS-T (5 minutes, room temperature, with gentle shaking; ~100 rpm). Then, the membranes were incubated with a secondary antibody (goat anti-rabbit IgG-HRP, Abcam) for 1-2 hours (room temperature, with gentle shaking; ~100 rpm). The secondary antibody was diluted in 15-20 mL Blotto. The membrane strips were then washed three times, each time in 20 mL TBST-T (5 minutes, room temperature, with gentle shaking; ~100 rpm). Protein bands on each membrane were detected via use of a chemiluminescence kit (EZ-Chemiluminescence Detection Kit for HRP, Geneflow) and a LAS-4000 imager (GE Healthcare). Where necessary, protein bands were quantified using AIDA Image Analyzer software (Raytest).

6.5 Transient expression assays and interaction studies

All centrifugation steps were carried out in at 4°C in a Sorvall Legend RT centrifuge fitted with 6441 swinging buckets.

6.5.1 Protoplast isolation

Protoplasts were isolated from the rosette leaves of 3-5-week-old plants via a modified version of an established protocol (Wu et al., 2009). Leaves were removed and adhered to electrical tape via their adaxial sides. Magic tape (Scotch) was then applied to the exposed abaxial sides, and the midvein of each leaf was crushed to promote the adherence of the leaf to the tape. Approximately 18-21 leaves were used in each extraction. The Magic tape was carefully removed, thus exposing the mesophyll cells via the removal of the abaxial epidermis. The leaves were then incubated in 25 mL cellulase solution (1% (w/v) cellulase 'Onozuka' R10 (Yakult), 0.25% (w/v) macerozyme 'Onozuka' R10 (Yakult), 0.4 M mannitol, 10 mM $CaCl_2$, 20 mM KCl, 0.1% (w/v) BSA, 20 mM MES pH

5.7) for 2-3 hours (room temperature, with gentle agitation; ~ 50 rpm). After incubation, the leaves were removed, and the protoplast suspension was transferred to a 50 mL Falcon tube (Corning) and stored on ice. The protoplasts were then very gently pelleted (60 × g, 5 minutes) and washed twice, each time in 25 mL ice-cold W5 solution (154 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, 5 mM glucose, 2 mM MES pH 5.7). Each wash consisted of the removal of the supernatant, the gentle resuspension of the pellet in fresh W5 solution, and centrifugation (40 × g, 4 minutes). After the second wash, the suspended protoplasts were incubated on ice for at least 30 minutes before being pelleted (40 × g, 4 minutes) and resuspended in 1-1.5 mL MMg solution (0.4 M mannitol, 15 mM $MgCl_2$, 4 mM MES pH 5.7). A typical protoplast isolation yielded 2-5 × 10^5 cells/mL.

6.5.2 Protoplast transfection

Protoplasts were transfected via use of an established method (Wu et al., 2009). Before starting, a polyethylene glycol (PEG) solution was prepared (40% (w/v) PEG 4000 (Fluka), 0.1 M $CaCl_2$, 0.2 M mannitol). In a standard transfection, 100 µL of protoplast suspension and 5 µg of plasmid DNA were transferred to a 2 mL microcentrifuge tube and very gently mixed. An aliquot of PEG solution (110 µL) was added to each sample, and the samples were then incubated at room temperature for 5 minutes. After the first incubation, 200 µL of room temperature W5 solution was added and the samples were again incubated at room temperature for 5 minutes. This step was repeated twice, first with 400 µL and then with 800 µL W5 solution. After the third incubation, the samples were gently centrifuged to pellet the protoplasts (40 × g, 3 minutes). Each pellet was then washed twice: in each case, the pellet was resuspended in 1.5 mL W5 solution and gently centrifuged (40 × g, 4 minutes). After the second wash, each pellet was resuspended in 600 µL W5 solution and transferred to the well of a culture plate (24-well, Corning). The culture plates were wrapped in aluminium foil and placed within small plastic boxes to exclude light and protect the samples. The boxes were incubated overnight at 20°C in a Percival growth chamber. The next day the protoplasts were pelleted by gentle

centrifugation (90-110 × g, 2 minutes). Most of the supernatant of each sample was removed, leaving approximately 50-60 µL of residual W5 solution. Each pellet was then resuspended in the remaining solution.

6.5.3 BiFC experiments

The coding sequence of each gene of interest (*SCE1*, *HPY2* and *SIZ1*) was cloned into a split-YFP vector (pSAT4(A)-nEYFP-N1) as described in detail in Section 6.3.3 (Tzfira et al., 2005). Genes encoding possible interaction partners and negative control proteins were cloned into a complementary vector (pSAT4(A)-cEYFP-N1) (Tzfira et al., 2005). The first vector tags its product with a fragment representing the N-terminal half of YFP, whereas the second tags its product with the C-terminal half of YFP. Each vector drives expression of the gene of interest using the constitutive cauliflower mosaic virus (CaMV) 35S promoter. The finished constructs were retrieved and grown up in 200 mL overnight cultures. The next day, plasmid DNA was isolated from each culture via use of a midiprep kit (Qiagen Plasmid Midi Kit). The manufacturer's instructions were followed with one minor exception: each pellet of plasmid DNA was eluted in a small volume (30-50 µL) of TE buffer to ensure that the final concentration was in excess of 1 µg/µL (as required by the protoplast transfection protocol).

Protoplasts were co-transfected with two constructs: one containing the SUMO system gene and the other containing a TOC complex or negative control gene. The protoplast transfection was carried out as described in Section 6.5.2; however, 5 µg plasmid DNA for each construct was used. The next day, protoplasts were harvested and the presence or absence of YFP signal was assessed via confocal microscopy (Leica TCS SP5). A robust YFP signal was taken to imply an interaction between the two expressed proteins. The images were analysed via use of the LAS X software package.

6.5.4 Immunoprecipitation

Each gene of interest was cloned into a YFP-tagging vector. In the case of HPY2, SCE1 and SIZ1, p2GWY7, a C-terminal tagging vector was used (Karimi et al., 2002). However, in the case of SUMO1, SUMO2 and SUMO3, p2YGW7, a highly similar, N-terminal tagging vector was used (Karimi et al., 2002); this was to prevent the proteolytic removal of the YFP tags from the nascent SUMO proteins, as each SUMO protein is synthesised with a short C-terminal extension that is removed prior to maturation (Benlloch and Lois, 2018). Each vector drives expression of the gene of interest using the constitutive CaMV 35S promoter. Plasmid DNA was isolated from 200mL LB cultures of each construct in exactly the same way as in the BiFC experiments.

Each construct was expressed in protoplasts using the same transfection protocol as in the BiFC experiments; however, the process was scaled up six-fold. For example, 30 µg plasmid DNA was used to transfect 600 µL protoplast suspension. After harvesting the protoplasts, each sample was resuspended in 500 µL lysis buffer (25 mM Tris-HCl pH 7.5, 150 mM NaCl, 10% (v/v) glycerol, 1 mM EDTA, 1% (v/v) Triton X-100, 5 µM bortezomib, 0.5% (v/v) PIC) and rotated for 30 minutes at 4°C. The samples were then cleared via centrifugation in a microcentrifuge (16,000 × *g*, 10 minutes, room temperature); the supernatants were recovered and transferred to fresh 2 mL microcentrifuge tubes. An aliquot of each sample (30 µL) was then removed and retained; this served as a 'total lysate' (TL) control in each case.

Each YFP-conjugated protein was pulled down via the use of GFP-Trap MA magnetic beads (Chromotek). Prior to use, each aliquot of beads was equilibrated with ice-cold immunoprecipitation (IP) buffer (the IP buffer was lysis buffer that did not contain bortezomib). In each case, 500 µL IP buffer was added to 25 µL beads, and the sample was thoroughly vortexed. The beads were then magnetically separated from the solution via the use of a MagJET Separation Rack (Thermo Fisher) and the IP buffer was removed and discarded. This equilibration step was carried out three times in total. The lysed

samples were then added to the equilibrated beads and the samples were rotated at 4°C for 1-4 hours or overnight. Next, a wash buffer was prepared (IP buffer with the addition of 10% (v/v) Triton X-100). The beads were magnetically separated from each sample and the solution was removed and transferred to a fresh 1.5 mL microcentrifuge tube. In some experiments, an aliquot of this solution (30 µL) was retained as a 'flow through' (FT) sample; these samples were taken to determine the binding efficiency of the target protein to the beads (in successful experiments, little if any target protein could be detected in these samples when resolved by immunoblotting; this indicated that most of the protein had been captured by the beads). The beads were then washed four times: in each case, the beads were resuspended in 500 µL ice-cold wash buffer and rotated at 4°C for 5-10 minutes. After the wash steps, the bound proteins were eluted from the beads to yield the 'immunoprecipitate' (IP) samples. Each sample was boiled (95°C, 10 minutes) in 30-50 µL of DTT-containing 2× Laemmli buffer (see Section 6.4.1). The TL and (where applicable) FT samples were diluted 1:1 in 2× Laemmli buffer and boiled in parallel with the IP samples. Each set of samples was resolved by western blotting.

6.6 Transmission electron microscopy

The TEM samples were fixed and imaged in collaboration with Dr Errin Johnson at the Dunn School of Pathology, University of Oxford. A modified version of an established protocol was used (Hyman and Jarvis, 2011).

6.6.1 Sample fixation

Tissue samples were taken from the leaves of plants that had been grown under identical conditions for four weeks. The samples were taken from three different genotypes: wild type, *ppi1*, and *ppi1 sce1-4*. Small (2 mm diameter) discs of tissue were removed from non-vascular sections of rosette leaves via the use of a biopsy punch. Each disc was immediately suspended in a droplet of fixative solution (2.5% (v/v) glutaraldehyde, 4% (v/v) paraformaldehyde (PFA), 0.1 M sodium cacodylate buffer (SCB), pH 6.9) and cut in half. Each fragment was then placed in a small plastic basket and

transferred to a microwave. The fragments (from now on referred to as ‘tissue samples’) were fixed at 37°C for 10 minutes, and then subjected to five cycles of alternating power (20 W, 2 minutes, 0 W, 2 minutes). The tissue samples were then transferred to 2 mL microcentrifuge tubes filled with fixative solution, and the tubes were manually shaken to encourage the tissue samples to fall to the bottom of each tube. The tubes were incubated at room temperature for approximately five hours, and then incubated at 4°C overnight.

The next day, the tissue samples were retrieved and washed twice with 1 mL 0.1 M SCB pH 6.9 (5 minutes with rotation). The tissue samples were then two more times, each time in 1 mL SCB (first wash: 45 minutes with rotation; second wash: 30 minutes with rotation). The tissue samples were then washed in 1 mL 50 mM glycine SCB (30 minutes with rotation), and then washed in 1 mL SCB again (20 minutes with rotation).

Next, the tissue samples were subjected to secondary fixation. Each sample was suspended in 1 mL secondary fixative solution (1% (v/v) osmium tetroxide, 1.5% (w/v) potassium ferrocyanide, SCB) and incubated at 4°C (2 hours with robust shaking; 200 rpm). The tissue samples were subsequently incubated at room temperature (50 minutes with rotation).

Finally, the tissue samples were subjected to tertiary fixation. Each sample was suspended in 1 mL tertiary fixative solution (2% (w/v) uranyl acetate) and incubated at 4°C in the dark (2 hours with rotation). Each tissue sample was subsequently washed three times, each time with 1 mL dH_2O (15 minutes with rotation).

6.6.2 Sample dehydration

Continuing to work in 2 mL microcentrifuge tubes, the fixed tissue samples were dehydrated via incubation with a progressively more concentrated series of ethanol solutions. In the first four steps, the samples were incubated at room temperature in 10%, 20%, 30% and 40% (v/v) ethanol solutions (in each case, 1 mL solution, and 10-minute

incubation with rotation). The tissue samples were then suspended sequentially in a 50% (v/v) ethanol solution for 20 minutes, a 60% (v/v) ethanol solution for 10 minutes, and a 70% (v/v) ethanol solution overnight (each step using 1 mL solution at room temperature). The next day, the tissue samples were suspended sequentially in ice-cold 80%, 90% and 100% (v/v) ethanol solutions and incubated at 4°C (30 minutes with rotation) in each case. Finally, the samples were transferred to 1 mL 100% anhydrous ethanol and incubated at 4°C (30 minutes with rotation).

6.6.3 Epoxy resin infiltration

The fixed tissue samples were incubated in a progressively concentrated series of resin solutions: 25%, 50% and 75% (v/v) epoxy resin (Agar Scientific) dissolved in acetone. All incubation steps were carried out in 1 mL volumes of solution in 2 mL microcentrifuge tubes at room temperature with rotation. The first incubation was allowed to proceed for 2 hours, the second for 3 hours, and the third for 1 hour. The tissue samples were centrifuged in a microcentrifuge (1 minute, 4000 × *g*) after each resin change. The tissue samples were then transferred to 100% resin rotated at room temperature overnight.

The next day, the resin was replaced with fresh resin twice (once at 12 pm, and once at 5 pm). The samples were left to incubate overnight. The next day at 10.30 am, the resin was replaced a for a final time.

6.6.4 Sample embedding and sectioning

First, embedding moulds were retrieved and prepared (TAAB Laboratories Equipment). Each mould had a small (1 mm deep) layer of resin added to its base and was incubated at 60°C to induce the resin to polymerise (2 hours). Next, the fixed tissue samples were added to the embedding moulds and overlaid with resin (one tissue sample per embedding mould). The blocks were then polymerised at 60°C for 48 hours. After polymerisation, thin sections of each block (<100 nm) were removed via use of an

ultramicrotome. These were then mounted onto tabbed support grids (Agar Scientific) and brought forward for imaging.

6.6.5 Imaging

The sections were imaged via use of a TEM microscope (FEI Tecnai T12). Three leaf fragments representing three biological replicates were fixed and sectioned per genotype, and ten TEM images were captured for each sample. The magnification level was held constant across the samples. Each image was analysed in ImageJ; the plan area of the chloroplasts contained within each image was measured.

6.7 The mutant screen

6.7.1 EMS mutagenesis

Seeds were mutagenized in accordance with a published protocol (Leyser and Furner, 1992). The *fts1* (*ppi2-3*) mutant used in this screen has already been described (Woodson et al., 2015; Ling et al., 2019). Approximately 29,000 *ppi2-3* seeds were exposed to ethyl methanesulfonate (EMS), a chemical mutagen. For safety reasons, the mutagenesis was carried out in an Atmos bag (Sigma). The seeds were contained within two modified 50 mL syringes. The tip of each syringe was removed and replaced with a layer of fine nylon mesh. This allowed the EMS solution to be easily drawn onto the seeds, and it also allowed the seeds to be agitated every hour to ensure even exposure to the solution (the EMS solution was ejected and replaced in each instance). First, the seeds were incubated overnight in a 0.1% (w/v) potassium chloride solution. Next, the seeds were soaked in an EMS solution (0.1 M sodium phosphate, 5% (v/v) dimethyl sulphoxide (DMSO), 100 mM EMS, pH 5) for three hours. After mutagenesis, the seeds were washed in 100 mM sodium thiosulphate twice (15 minutes) to inactivate the EMS. Next, the seeds were washed in distilled water twice (15 minutes). The seeds were then removed from the Atmos bag, transferred to filter paper, and dried out overnight.

The next day, the M1 seeds were sown out onto soil. The seeds were sown out in 100 separate batches. Each batch contained approximately 290 seeds grown at high density in a single, small rectangular pot (15.5 × 21 × 5.5 cm (length/width/height)). A standard module tray held two pots (A and B), and so the batches were named with reference to both the module tray and the pot (e.g. batch 10-A, batch 10-B, etc.). The plants were grown to maturity and allowed to set seed. The pooled M2 seed of each batch was separately harvested, and care was taken to prevent cross-contamination of seed between the batches.

6.7.2 M2 screening

The M2 seeds were screened for visible suppressor mutants over several months. Approximately 10% of each batch of seed was grown *in vitro* for 14 days, and, at this point, the plates were screened by eye for seedlings that appeared greener than their neighbours, and greener than the un-mutagenized *ppi2-3* progenitor line. Selected putative mutant seedlings were transferred to soil, grown to maturity, and allowed to set seed. Due to time constraints, only 34 of the 100 batches were screened. The full set of M3 suppressor mutants were then grown on soil alongside wild-type and *ppi2-3* control plants. After approximately four weeks, the plants were photographed and the concentration of chlorophyll in each mutant was measured.

6.8 Tables

Table 6.1: Primers used for genotyping purposes.

Primer name	Sequence	Used to genotype…
SUMO1 F(2)	AAAAAGCAGGCTCCACAAAAGCCACGGCCAATTAG	*sum1-1*
SUMO1 R(2)	AGAAAGCTGGGTTCCATTCATATCACACACAAGCCC	*sum1-1*
SUMO2 F	CGTTGTTGGTACTTGGTTGG	*sum2-1*
SUMO2 R	CAAAACTCTAAACTGGTCGG	*sum2-1*
SUMO3 F	ACAGACTGGAGTTTTTGTTTC	*sum3-1*
SUMO3 R	CTCATGAGTCATTTACACACACG	*sum3-1*
SCE1 F	CGCCGCGAAATCTGGACC	*sce1-4*
SCE1 R	TTCCTCTCTTCAGCTAAACG	*sce1-4*
SIZ1 F(2)	GCAAACAGGGAAAGAAGCAGG	*siz1-4*
SIZ1 R(2)	CATTGAGTCTGTTTCTAGCG	*siz1-4*
LBb1	GCGTGGACCGCTTGCTGCAACT	SALK lines (left border)
LB1	GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC	SAIL lines (left border)

For a more detailed description of the T-DNA mutants used in this study, please refer to Chapter 2.

Table 6.2: Primers used in restriction cloning procedures.

Primer name	Sequence	Used to generate…
SCE1 F (HindIII)	AAAAGCTTATGGCTAGTGGAATCGCTC	pSAT4(A)-nEYFP-N1 SCE1
SCE1 R (EcoRI)	AAGAATTCGACAAGAGCAGGATACTGCTTG	pSAT4(A)-nEYFP-N1 SCE1
SIZ1 F (HindIII)	AAAAGCTTATGGATTTGGAAGCTAATTGTAAG	pSAT4(A)-nEYFP-N1 SIZ1
SIZ1 R (EcoRI)	AAGAATTCAACTCCGGTGTCTTGTCTGATG	pSAT4(A)-nEYFP-N1 SIZ1
HPY2 (HindIII)	AAAAGCTTATGGCGTCGGCGTCCTC	pSAT4(A)-nEYFP-N1 HPY2
HPY2 (SalI)	AAGTCGACGATCTTCATCCACATCTTCTGTG	pSAT4(A)-nEYFP-N1 HPY2

Restriction sequences are underlined.

Table 6.3: Primers used in Gateway cloning procedures.

Primer name	Sequence	Used to generate…
SCE1 Gateway F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGGCTAGTGGAATCGCTC	p2GWY7 SCE1
SCE1 Gateway R	GGGGACCACTTTGTACAAGAAAGCTGGGTTGACAAGAGCAGGATACTGC	p2GWY7 SCE1
SIZ1 Gateway F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGGATTTGGAAGCTAATTGTAAG	p2GWY7 SIZ1
SIZ1 Gateway R	GGGGACCACTTTGTACAAGAAAGCTGGGTTAACTCCGGTGTCTTGTCTGATG	p2GWY7 SIZ1
HPY2 Gateway F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGGCGTCGGCGTCCTCG	p2GWY7 HPY2
HPY2 Gateway R	GGGGACCACTTTGTACAAGAAAGCTGGGTTATCTTCATCCACATCTTCTGTG	p2GWY7 HPY2
SUMO1 Gateway F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCTCTGCAAACCAGGAGGAAGACAAG	p2YGW7 SUMO1
SUMO1 Gateway R	GGGGACCACTTTGTACAAGAAAGCTGGGTTTCAGGCCGTAGCACCACC	p2YGWY SUMO1
SUMO2 Gateway F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCTCTGCTACTCCGGAAGAAGAC	p2YGW7 SUMO2
SUMO2 Gateway R	GGGGACCACTTTGTACAAGAAAGCTGGGTTCTAAAAGCAGAAGAGCTTCAGGCC	p2YGW7 SUMO2
SUMO3 Gateway F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCTCTAACCCTCAAGATGACAAGCCC	p2YGW7 SUMO3
SUMO3 Gateway R	GGGGACCACTTTGTACAAGAAAGCTGGGTTTTAAAGCCCATTATGATCGAAAAGC	p2YGW7 SUMO3

AttB1 and AttB2 sequences are underlined.

Table 6.4: Primers used in RT-PCR experiments.

Primer name	Sequence	Used to amplify…
SUMO1 F(2)	AAAAAGCAGGCTCCACAAAAGCCACGGCCAATTAG	*SUMO1*
SUMO1 R(2)	AGAAAGCTGGGTTCCATTCATATCACACACAAGCCC	*SUMO1*
SUMO3 F	ACAGACTGGAGTTTTTGTTTC	*SUMO3*
SUMO3 R	CTCATGAGTCATTTACACACACG	*SUMO3*
eIF4E1 F	AAGATTTGAGAGGTTTCAAGCGGTGTAAG	*eIF4E1*
eIF4E1 R	AAACAATGGCGGTAGAAGACACTC	*eIF4E1*

7 Bibliography

Agne, B., Infanger, S., Wang, F., Hofstetter, V., Rahim, G., Martin, M., Lee, D.W., Hwang, I., Schnell, D., and Kessler, F. (2009). A toc159 import receptor mutant, defective in hydrolysis of GTP, supports preprotein import into chloroplasts. Journal of Biological Chemistry **284,** 8670-8679.

Agne, B., Andres, C., Montandon, C., Christ, B., Ertan, A., Jung, F., Infanger, S., Bischof, S., Baginsky, S., and Kessler, F. (2010). The acidic A-domain of *Arabidopsis* TOC159 occurs as a hyperphosphorylated protein. Plant Physiology **153,** 1016-1030.

Ahner, A., Gong, X., Schmidt, B.Z., Peters, K.W., Rabeh, W.M., Thibodeau, P.H., Lukacs, G.L., and Frizzell, R.A. (2013). Small heat shock proteins target mutant cystic fibrosis transmembrane conductance regulator for degradation via a small ubiquitin-like modifier-dependent pathway. Molecular Biology of the Cell **24,** 74-84.

Akita, M., Nielsen, E., and Keegstra, K. (1997). Identification of protein transport complexes in the chloroplastic envelope membranes via chemical cross-linking. Journal of Cell Biology **136,** 983-994.

Akutsu, M., Dikic, I., and Bremm, A. (2016). Ubiquitin chain diversity at a glance. Journal of Cell Science **129,** 875-880.

Amerik, A.Y., and Hochstrasser, M. (2004). Mechanism and function of deubiquitinating enzymes. Biochimica et Biophysica Acta **1695,** 189-207.

Andres, C., Agne, B., and Kessler, F. (2010). The TOC complex: preprotein gateway to the chloroplast. Biochimica et Biophysica Acta **1803,** 715-723.

Aronsson, H., Boij, P., Patel, R., Wardle, A., Topel, M., and Jarvis, P. (2007). Toc64/OEP64 is not essential for the efficient import of proteins into chloroplasts in *Arabidopsis thaliana*. Plant Journal **52,** 53-68.

Baek, G.H., Cheng, H., Choe, V., Bao, X., Shao, J., Luo, S., and Rao, H. (2013). Cdc48: a swiss army knife of cell biology. Journal of Amino Acids **2013,** 183421.

Bailey, M., Srivastava, A., Conti, L., Nelis, S., Zhang, C.J., Florance, H., Love, A., Milner, J., Napier, R., Grant, M., and Sadanandom, A. (2016). Stability of small ubiquitin-like modifier (SUMO) proteases OVERLY TOLERANT TO SALT1 and-2 modulates salicylic acid signalling and SUMO1/2 conjugation in *Arabidopsis thaliana*. Journal of Experimental Botany **67,** 353-363.

Baldwin, A., Wardle, A., Patel, R., Dudley, P., Park, S.K., Twell, D., Inoue, K., and Jarvis, P. (2005). A molecular-genetic study of the *Arabidopsis* Toc75 gene family. Plant Physiology **138,** 715-733.

Balsera, M., Soll, J., and Buchanan, B. (2010). Redox extends its regulatory reach to chloroplast protein import. Trends in Plant Science **15,** 515-521.

Bauer, J., Chen, K., Hiltbunner, A., Wehrli, E., Eugster, M., Schnell, D., and Kessler, F. (2000). The major protein import receptor of plastids is essential for chloroplast biogenesis. Nature **403,** 203-207.

Bedard, J., Kubis, S., Bimanadham, S., and Jarvis, P. (2007). Functional similarity between the chloroplast translocon component, Tic40, and the human co-chaperone, Hsp70-interacting protein (Hip). Journal of Biological Chemistry **282,** 21404-21414.

Benlloch, R., and Lois, L.M. (2018). Sumoylation in plants: mechanistic insights and its role in drought stress. Journal of Experimental Botany **69,** 4539-4554.

Bergink, S., Ammon, T., Kern, M., Schermelleh, L., Leonhardt, H., and Jentsch, S. (2013). Role of Cdc48/p97 as a SUMO-targeted segregase curbing Rad51-Rad52 interaction. Nature Cell Biology **15,** 526-532.

Bolter, B., May, T., and Soll, J. (1998). A protein import receptor in pea chloroplasts, Toc86, is only a proteolytic fragment of a larger polypeptide. FEBS Letters **441,** 59-62.

Bragoszewski, P., Turek, M., and Chacinska, A. (2017). Control of mitochondrial biogenesis and function by the ubiquitin - proteasome system. Open Biology **7**.

Budhiraja, R., Hermkes, R., Muller, S., Schmidt, J., Colby, T., Panigrahi, K., Coupland, G., and Bachmair, A. (2009). Substrates related to chromatin and to RNA-dependent processes are modified by *Arabidopsis* SUMO isoforms that differ in a conserved residue with influence on desumoylation. Plant Physiology **149,** 1529-1540.

Butterfield, N.J., Knoll, A.H., and Swett, K. (1990). A bangiophyte red alga from the Proterozoic of arctic Canada. Science **250,** 104-107.

Caliebe, A., Grimm, R., Kaiser, G., Lübeck, J., Soll, J., and Heins, L. (1997). The chloroplastic protein import machinery contains a Rieske-type iron–sulfur cluster and a mononuclear iron-binding protein. The EMBO Journal **16,** 7342-7350.

Castano-Miquel, L., Segui, J., Manrique, S., Teixeira, I., Carretero-Paulet, L., Atencio, F., and Lois, L.M. (2013). Diversification of SUMO-activating enzyme in *Arabidopsis*: Implications in SUMO conjugation. Molecular Plant **6,** 1646-1660.

Castro, P.H., Bachmair, A., Bejarano, E.R., Coupland, G., Lois, L.M., Sadanandom, A., van den Burg, H.A., Vierstra, R.D., and Azevedo, H. (2018). Revised nomenclature and functional overview of the ULP gene family of plant deSUMOylating proteases. Journal of Experimental Botany **69,** 4505-4509.

Chen, K., Chen, X., and Schnell, D.J. (2000). Initial binding of preproteins involving the Toc159 receptor can be bypassed during protein import into chloroplasts. Plant Physiology **122,** 813-822.

Chen, X.J., Smith, M.D., Fitzpatrick, L., and Schnell, D.J. (2002). In vivo analysis of the role of atTic20 in protein import into chloroplasts. Plant Cell **14,** 641-654.

Chen, Y.L., Chen, L.J., Chu, C.C., Huang, P.K., Wen, J.R., and Li, H.M. (2018). TIC236 links the outer and inner membrane translocons of the chloroplast. Nature **564,** 125-129.

Chou, M.L., Fitzpatrick, L.M., Tu, S.L., Budziszewski, G., Potter-Lewis, S., Akita, M., Levin, J.Z., Keegstra, K., and Li, H.M. (2003). Tic40, a membrane-anchored co-chaperone homolog in the chloroplast protein translocon. The EMBO Journal **22,** 2970-2980.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant Journal **16,** 735-743.

Constan, D., Patel, R., Keegstra, K., and Jarvis, P. (2004). An outer envelope membrane component of the plastid protein import apparatus plays an essential role in *Arabidopsis*. Plant Journal **38,** 93-106.

Conti, L., Price, G., O'Donnell, E., Schwessinger, B., Dominy, P., and Sadanandom, A. (2008). Small ubiquitin-like modifier proteases OVERLY TOLERANT TO SALT1 and-2 regulate salt stress responses in *Arabidopsis*. Plant Cell **20,** 2894-2908.

Daviere, J.M., and Achard, P. (2013). Gibberellin signaling in plants. Development **140,** 1147-1151.

de Vries, J., Sousa, F.L., Bolter, B., Soll, J., and Gould, S.B. (2015). YCF1: A Green TIC? Plant Cell **27,** 1827-1833.

Depaux, A., Regnier-Ricard, F., Germani, A., and Varin-Blank, N. (2007). A crosstalk between hSiah2 and Pias E3-ligases modulates Pias-dependent activation. Oncogene **26,** 6665-6676.

Deshaies, R.J., and Joazeiro, C.A. (2009). RING domain E3 ubiquitin ligases. Annual Review of Biochemistry **78,** 399-434.

Duy, D., Wanner, G., Meda, A.R., von Wiren, N., Soll, J., and Philippar, K. (2007). PIC1, an ancient permease in *Arabidopsis* chloroplasts, mediates iron transport. Plant Cell **19,** 986-1006.

Elrouby, N. (2015). Analysis of small ubiquitin-Like modifier (SUMO) targets reflects the essential nature of protein SUMOylation and provides insight to elucidate the role of SUMO in plant development. Plant Physiology **169,** 1006-1017.

Elrouby, N., and Coupland, G. (2010). Proteome-wide screens for small ubiquitin-like modifier (SUMO) substrates identify *Arabidopsis* proteins implicated in diverse biological processes. Proceedings of the National Academy of Sciences of the United States of America **107,** 17415-17420.

Flores-Perez, U., and Jarvis, P. (2013). Molecular chaperone involvement in chloroplast protein import. Biochimica et Biophysica Acta **1833,** 332-340.

Flores-Perez, U., Bedard, J., Tanabe, N., Lymperopoulos, P., Clarke, A.K., and Jarvis, P. (2016). Functional analysis of the Hsp93/ClpC chaperone at the chloroplast envelope. Plant Physiology **170,** 147-162.

Frydman, J., and Hohfeld, J. (1997). Chaperones get in touch: The hip-hop connection. Trends in Biochemical Sciences **22,** 87-92.

Gentle, I.E., Burri, L., and Lithgow, T. (2005). Molecular architecture and function of the Omp85 family of proteins. Molecular Microbiology **58,** 1216-1225.

Gutensohn, M., Schulz, B., Nicolay, P., and Flugge, U.I. (2000). Functional analysis of the two *Arabidopsis* homologues of Toc34, a component of the chloroplast protein import apparatus. Plant Journal **23,** 771-783.

Han, T.M., and Runnegar, B. (1992). Megascopic eukaryotic algae from the 2.1-billion-year-old negaunee iron-formation, Michigan. Science **257,** 232-235.

Harder, Z., Zunino, R., and McBride, H. (2004). Sumo1 conjugates mitochondrial substrates and participates in mitochondrial fission. Current Biology **14,** 340-345.

Heins, L., Mehrle, A., Hemmler, R., Wagner, R., Kuchler, M., Hormann, F., Sveshnikov, D., and Soll, J. (2002). The preprotein conducting channel at the inner envelope membrane of plastids. The EMBO Journal **21,** 2616-2625.

Hermkes, R., Fu, Y.F., Nurrenberg, K., Budhiraja, R., Schmelzer, E., Elrouby, N., Dohmen, R.J., Bachmair, A., and Coupland, G. (2011). Distinct roles for *Arabidopsis* SUMO protease ESD4 and its closest homolog ELS1. Planta **233,** 63-73.

Hiltbrunner, A., Grunig, K., Alvarez-Huerta, M., Infanger, S., Bauer, J., and Kessler, F. (2004). AtToc90, a new GTP-binding component of the *Arabidopsis* chloroplast protein import machinery. Plant Molecular Biology **54,** 427-440.

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

Hirabayashi, Y., Kikuchi, S., Oishi, M., and Nakai, M. (2011). In vivo studies on the roles of two closely related *Arabidopsis* Tic20 proteins, AtTic20-I and AtTic20-IV. Plant Cell Physiology **52,** 469-478.

Hirsch, S., Muckel, E., Heemeyer, F., von Heijne, G., and Soll, J. (1994). A receptor component of the chloroplast protein translocation machinery. Science **266,** 1989-1992.

Hofgen, R., and Willmitzer, L. (1988). Storage of competent cells for *Agrobacterium* transformation. Nucleic Acids Research **16,** 9877.

Hofmann, N.R., and Theg, S.M. (2005). Toc64 is not required for import of proteins into chloroplasts in the moss *Physcomitrella patens*. Plant Journal **43,** 675-687.

Hormann, F., Kuchler, M., Sveshnikov, D., Oppermann, U., Li, Y., and Soll, J. (2004). Tic32, an essential component in chloroplast biogenesis. Journal of Biological Chemistry **279,** 34756-34762.

Huang, L.X., Yang, S.G., Zhang, S.C., Liu, M., Lai, J.B., Qi, Y.L., Shi, S.F., Wang, J.X., Wang, Y.Q., Xie, Q., and Yang, C.W. (2009). The *Arabidopsis* SUMO E3 ligase AtMMS21, a homologue of NSE2/MMS21, regulates cell proliferation in the root. Plant Journal **60,** 666-678.

Huang, W.H., Ling, Q.H., Bedard, J., Lilley, K., and Jarvis, P. (2011). *In vivo* analyses of the roles of essential Omp85-related proteins in the chloroplast outer envelope membrane. Plant Physiology **157,** 147-159.

Hyman, S., and Jarvis, R.P. (2011). Studying *Arabidopsis* chloroplast structural organisation using transmission electron microscopy. Methods in Molecular Biology **774,** 113-132.

Inaba, T., Li, M., Alvarez-Huerta, M., Kessler, F., and Schnell, D.J. (2003). AtTic110 functions as a scaffold for coordinating the stromal events of protein import into chloroplasts. Journal of Biological Chemistry **278,** 38617-38627.

Infanger, S., Bischof, S., Hiltbrunner, A., Agne, B., Baginsky, S., and Kessler, F. (2011). The chloroplast import receptor Toc90 partially restores the accumulation of Toc159 client proteins in the *Arabidopsis thaliana ppi2* mutant. Molecular Plant **4,** 252-263.

Inoue, H., Rounds, C., and Schnell, D.J. (2010). The molecular basis for distinct pathways for protein import into *Arabidopsis* chloroplasts. Plant Cell **22,** 1947-1960.

Ishida, T., Yoshimura, M., Miura, K., and Sugimoto, K. (2012). MMS21/HPY2 and SIZ1, two *Arabidopsis* SUMO E3 ligases, have distinct functions in development. Plos One **7**(10): e46897.

Ivanova, Y., Smith, M.D., Chen, K., and Schnell, D.J. (2004). Members of the Toc159 import receptor family represent distinct pathways for protein targeting to plastids. Molecular Biology of the Cell **15,** 3379-3392.

Jackson-Constan, D., and Keegstra, K. (2001). *Arabidopsis* genes encoding components of the chloroplastic protein import apparatus. Plant Physiology **125,** 1567-1576.

Jackson, D.T., Froehlich, J.E., and Keegstra, K. (1998). The hydrophilic domain of Tic110, an inner envelope membrane component of the chloroplastic protein translocation apparatus, faces the stromal compartment. Journal of Biological Chemistry **273,** 16583-16588.

Jarvis, P. (2008). Targeting of nucleus-encoded proteins to chloroplasts in plants. New Phytologist **179,** 257-285.

Jarvis, P., and Lopez-Juez, E. (2013). Biogenesis and homeostasis of chloroplasts and other plastids. Nature Reviews Molecular Cell Biology **14,** 787-802.

Jarvis, P., Chen, L.J., Li, H.M., Pete, C.A., Fankhauser, C., and Chory, J. (1998). An *Arabidopsis* mutant defective in the plastid general protein import apparatus. Science **282,** 100-103.

Jin, J.B., Jin, Y.H., Lee, J., Miura, K., Yoo, C.Y., Kim, W.Y., Van Oosten, M., Hyun, Y., Somers, D.E., Lee, I., Yun, D.J., Bressan, R.A., and Hasegawa, P.M. (2008). The SUMO E3 ligase, AtSIZ1, regulates flowering by controlling a salicylic acid-mediated floral promotion pathway and through affects on FLC chromatin structure. Plant Journal **53,** 530-540.

Karimi, M., Inze, D., and Depicker, A. (2002). GATEWAY vectors for Agrobacterium-mediated plant transformation. Trends in Plant Science **7,** 193-195.

Kasmati, A.R., Töpel, M., Patel, R., Murtaza, G., and Jarvis, P. (2011). Molecular and genetic analyses of Tic20 homologues in *Arabidopsis thaliana* chloroplasts. Plant Journal **66,** 877-889.

Kasmati, A.R., Topel, M., Khan, N.Z., Patel, R., Ling, Q.H., Karim, S., Aronsson, H., and Jarvis, P. (2013). Evolutionary, molecular and genetic snalyses of Tic22 homologues in *Arabidopsis thaliana* chloroplasts. Plos One **8**(5):e63863.

Kessler, F., and Blobel, G. (1996). Interaction of the protein import and folding machineries of the chloroplast. Proceedings of the National Academy of Sciences of the United States of America **93,** 7684-7699.

Kessler, F., Blobel, G., Patel, H.A., and Schnell, D.J. (1994). Identification of two GTP-binding proteins in the chloroplast protein import machinery. Science **266,** 1035-1039.

Kikuchi, S., Hirohashi, T., and Nakai, M. (2006). Characterization of the preprotein translocon at the outer envelope membrane of chloroplasts by blue native PAGE. Plant Cell Physiology **47,** 363-371.

Kikuchi, S., Bedard, J., Hirano, M., Hirabayashi, Y., Oishi, M., Imai, M., Takase, M., Ide, T., and Nakai, M. (2013). Uncovering the protein translocon at the chloroplast inner envelope membrane. Science **339,** 571-574.

Kim, S.I., Kwak, J.S., Song, J.T., and Seo, H.S. (2016). The E3 SUMO ligase AtSIZ1 functions in seed germination in *Arabidopsis*. Physiologia Plantarum **158,** 256-271.

Koegl, M., Hoppe, T., Schlenker, S., Ulrich, H.D., Mayer, T.U., and Jentsch, S. (1999). A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. Cell **96,** 635-644.

Kohler, J.B., Tammsalu, T., Jorgensen, M.M., Steen, N., Hay, R.T., and Thon, G. (2015). Targeting of SUMO substrates to a Cdc48-Ufd1-Npl4 segregase and STUbL pathway in fission yeast. Nature Communications **6,** 8827.

Komander, D., and Rape, M. (2012). The ubiquitin code. Annual Review of Biochemistry **81,** 203-229.

Kouranov, A., and Schnell, D.J. (1997). Analysis of the interactions of preproteins with the import machinery over the course of protein import into chloroplasts. Journal of Cell Biology **139,** 1677-1685.

Kouranov, A., Chen, X.J., Fuks, B., and Schnell, D.J. (1998). Tic20 and Tic22 are new components of the protein import apparatus at the chloroplast inner envelope membrane. Journal of Cell Biology **143,** 991-1002.

Kovacheva, S., Bedard, J., Patel, R., Dudley, P., Twell, D., Rios, G., Koncz, C., and Jarvis, P. (2005). *In vivo* studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. Plant Journal **41,** 412-428.

Kovacs-Bogdan, E., Benz, J.P., Soll, J., and Bolter, B. (2011). Tic20 forms a channel independent of Tic110 in chloroplasts. BMC Plant Biology **11,** 133.

Kubis, S., Baldwin, A., Patel, R., Razzaq, A., Dupree, P., Lilley, K., Kurth, J., Leister, D., and Jarvis, P. (2003). The *Arabidopsis ppi1* mutant is specifically defective in the expression, chloroplast import, and accumulation of photosynthetic proteins. Plant Cell **15,** 1859-1871.

Kubis, S., Patel, R., Combe, J., Bedard, J., Kovacheva, S., Lilley, K., Biehl, A., Leister, D., Rios, G., Koncz, C., and Jarvis, P. (2004). Functional specialization amongst the *Arabidopsis* Toc159 family of chloroplast protein import receptors. Plant Cell **16,** 2059-2077.

Kurepa, J., Walker, J.M., Smalle, J., Gosink, M.M., Davis, S.J., Durham, T.L., Sung, D.Y., and Vierstra, R.D. (2003). The small ubiquitin-like modifier (SUMO) protein modification system in *Arabidopsis*. Accumulation of SUMO1 and -2 conjugates is increased by stress. Journal of Biological Chemistry **278,** 6862-6872.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227,** 680-685.

Lee, J., Nam, J., Park, H.C., Na, G., Miura, K., Jin, J.B., Yoo, C.Y., Baek, D., Kim, D.H., Jeong, J.C., Kim, D., Lee, S.Y., Salt, D.E., Mengiste, T., Gong, Q.Q., Ma, S.S., Bohnert, H.J., Kwak, S.S., Bressan, R.A., Hasegawa, P.M., and Yun, D.J. (2007). Salicylic acid-mediated innate immunity in *Arabidopsis* is regulated by SIZ1 SUMO E3 ligase. Plant Journal **49,** 79-90.

Lee, K.H., Kim, S.J., Lee, Y.J., Jin, J.B., and Hwang, I. (2003). The M domain of atToc159 plays an essential role in the import of proteins into chloroplasts and chloroplast biogenesis. Journal of Biological Chemistry **278,** 36794-36805.

Lee, Y.J., Kim, D.H., Kim, Y.W., and Hwang, I. (2001). Identification of a signal that distinguishes between the chloroplast outer envelope membrane and the endomembrane system *in vivo*. Plant Cell **13**, 2175-2190.

Leyser, H., and Furner, I. (1992). EMS Mutagenesis of *Arabidopsis*. *Arabidopsis*: The Compleat Guide.

Lin, X.L., Niu, D., Hu, Z.L., Kim, D.H., Jin, Y.H., Cai, B., Liu, P., Miura, K., Yun, D.J., Kim, W.Y., Lin, R.C., and Jin, J.B. (2016). An *Arabidopsis* SUMO E3 ligase, SIZ1, negatively regulates photomorphogenesis by promoting COP1 activity. Plos Genetics **12**(4): e1006016.

Ling, Q., and Jarvis, P. (2015). Regulation of chloroplast protein import by the ubiquitin E3 ligase SP1 is important for stress tolerance in plants. Current Biology **25,** 2527-2534.

Ling, Q., Huang, W., and Jarvis, P. (2011). Use of a SPAD-502 meter to measure leaf chlorophyll concentration in *Arabidopsis thaliana*. Photosynthesis Research **107,** 209-214.

Ling, Q., Li, N., and Jarvis, P. (2017). Chloroplast ubiquitin E3 ligase SP1: Does it really function in peroxisomes? Plant Physiology **175,** 586-588.

Ling, Q., Broad, W., Trosch, R., Topel, M., Demiral Sert, T., Lymperopoulos, P., Baldwin, A., and Jarvis, R.P. (2019). Ubiquitin-dependent chloroplast-associated protein degradation in plants. Science **363**.

Ling, Q.H., Huang, W.H., Baldwin, A., and Jarvis, P. (2012). Chloroplast biogenesis is regulated by direct action of the ubiquitin-proteasome system. Science **338,** 655-659.

Liu, C., Yu, H., and Li, L. (2019). SUMO modification of LBD30 by SIZ1 regulates secondary cell wall formation in *Arabidopsis thaliana*. Plos Genetics **15**(1): e1007928.

Liu, L.P., Jiang, Y., Zhang, X.M., Wang, X., Wang, Y.B., Han, Y.Z., Coupland, G., Jin, J.B., Searle, I., Fu, Y.F., and Chen, F.L. (2017). Two SUMO proteases SUMO PROTEASE RELATED TO FERTILITY1 and 2 are required for fertility in *Arabidopsis*. Plant Physiology **175,** 1703-1719.

Liu, M., Shi, S.F., Zhang, S.C., Xu, P.L., Lai, J.B., Liu, Y.Y., Yuan, D.K., Wang, Y.Q., Du, J.J., and Yang, C.W. (2014). SUMO E3 ligase AtMMS21 is required for normal meiosis and gametophyte development in *Arabidopsis*. BMC Plant Biology **14**, 153.

Liu, Y., Lai, J., Yu, M., Wang, F., Zhang, J., Jiang, J., Hu, H., Wu, Q., Lu, G., Xu, P., and Yang, C. (2016). The *Arabidopsis* SUMO E3 ligase AtMMS21 dissociates the E2Fa/DPa complex in cell cycle regulation. Plant Cell **28,** 2225-2237.

Lopez-Juez, E., and Pyke, K.A. (2005). Plastids unleashed: their development and their integration in plant development. The International Journal of Developmental Biology **49,** 557-577.

Marin, B., Nowack, E.C., and Melkonian, M. (2005). A plastid in the making: evidence for a second primary endosymbiosis. Protist **156,** 425-432.

Menzies, S.A., Volkmar, N., van den Boomen, D.J., Timms, R.T., Dickson, A.S., Nathan, J.A., and Lehner, P.J. (2018). The sterol-responsive RNF145 E3 ubiquitin ligase mediates the degradation of HMG-CoA reductase together with gp78 and Hrd1. Elife **7**: e40009.

Miura, K., Lee, J., Jin, J.B., Yoo, C.Y., Miura, T., and Hasegawa, P.M. (2009). Sumoylation of ABI5 by the *Arabidopsis* SUMO E3 ligase SIZ1 negatively regulates abscisic acid signaling. Proceedings of the National Academy of Sciences of the United States of America **106,** 5418-5423.

Miura, K., Jin, J.B., Lee, J., Yoo, C.Y., Stirm, V., Miura, T., Ashworth, E.N., Bressan, R.A., Yun, D.J., and Hasegawa, P.M. (2007). SIZ1-mediated sumoylation of ICE1 controls CBF3/DREB1A expression and freezing tolerance in *Arabidopsis*. Plant Cell **19,** 1403-1414.

Miura, K., Lee, J., Gong, Q.Q., Ma, S.S., Jin, J.B., Yoo, C.Y., Miura, T., Sato, A., Bohnert, H.J., and Hasegawa, P.M. (2011). SIZ1 regulation of phosphate starvation-induced root architecture remodeling involves the control of auxin accumulation. Plant Physiology **155,** 1000-1012.

Miura, K., Rus, A., Sharkhuu, A., Yokoi, S., Karthikeyan, A.S., Raghothama, K.G., Back, D., Koo, Y.D., Jin, J.B., Bressan, R.A., Yun, D.J., and Hasegawa, P.M. (2005). The *Arabidopsis* SUMO E3 ligase SIZ1 controls phosphate deficiency responses Proceedings of the National Academy of Sciences of the United States of America **102,** 9734-9734.

Murtas, G., Reeves, P.H., Fu, Y.F., Bancroft, I., Dean, C., and Coupland, G. (2003). A nuclear protease required for flowering-time regulation in *Arabidopsis* reduces the abundance of SMALL UBIQUITIN-RELATED MODIFIER conjugates. Plant Cell **15,** 2308-2319.

Nakai, M. (2015a). YCF1: A Green TIC: Response to the de Vries et al. Commentary. Plant Cell **27,** 1834-1838.

Nakai, M. (2015b). The TIC complex uncovered: The alternative view on the molecular mechanism of protein translocation across the inner envelope membrane of chloroplasts. Biochim Biophys Acta **1847,** 957-967.

Nie, M., Aslanian, A., Prudden, J., Heideker, J., Vashisht, A.A., Wohlschlegel, J.A., Yates, J.R., and Boddy, M.N. (2012). Dual recruitment of Cdc48 (p97)-Ufd1-Npl4 ubiquitin-selective segregase by small ubiquitin-like modifier protein (SUMO) and ubiquitin in SUMO-targeted ubiquitin ligase-mediated genome stability functions. Journal of Biological Chemistry **287,** 29610-29619.

Nielsen, E., Akita, M., DavilaAponte, J., and Keegstra, K. (1997). Stable association of chloroplastic precursors with protein translocation complexes that contain proteins from both envelope membranes and a stromal Hsp 100 molecular chaperone. The EMBO Journal **16,** 935-946.

Nishimura, K., Kato, Y., and Sakamoto, W. (2016). Chloroplast proteases: updates on proteolysis within and across suborganellar compartments. Plant Physiology **171,** 2280-2293.

Novatchkova, M., Tomanov, K., Hofmann, K., Stuible, H.P., and Bachmair, A. (2012). Update on sumoylation: defining core components of the plant SUMO conjugation system by phylogenetic comparison. New Phytologist **195,** 23-31.

O'Neil, P.K., Richardson, L.G.L., Paila, Y.D., Piszczek, G., Chakravarthy, S., Noinaj, N., and Schnell, D. (2017). The POTRA domains of Toc75 exhibit chaperone-like function to facilitate import into chloroplasts. Proceedings of the National Academy of Sciences of the United States of America **114,** E4868-E4876.

Paila, Y.D., Richardson, L.G., Inoue, H., Parks, E.S., McMahon, J., Inoue, K., and Schnell, D.J. (2016). Multi-functional roles for the polypeptide transport associated domains of Toc75 in chloroplast protein import. Elife **5**: e12631.

Park, B.S., Song, J.T., and Seo, H.S. (2011a). *Arabidopsis* nitrate reductase activity is stimulated by the E3 SUMO ligase AtSIZ1. Nature Communications **2**:400.

Park, H.J., Kim, W.Y., Park, H.C., Lee, S.Y., Bohnert, H.J., and Yun, D.J. (2011b). SUMO and SUMOylation in plants. Molecular Cell **32,** 305-316.

Perry, S.E., and Keegstra, K. (1994). Envelope membrane proteins that interact with chloroplastic precursor proteins. Plant Cell **6,** 93-105.

Pogson, B.J., Ganguly, D., and Albrecht-Borth, V. (2015). Insights into chloroplast biogenesis and development. Biochimica Et Biophysica Acta-Bioenergetics **1847,** 1017-1024.

Qbadou, S., Becker, T., Mirus, O., Tews, I., Soll, J., and Schleiff, E. (2006). The molecular chaperone Hsp90 delivers precursor proteins to the chloroplast import receptor Toc64. The EMBO Journal **25,** 1836-1847.

Rodriguez, C.M., Freire, M.A., Camilleri, C., and Robaglia, C. (1998). The *Arabidopsis thaliana* cDNAs coding for eIF4E and eIF(iso)4E are not functionally equivalent for yeast complementation and are differentially expressed during plant development. Plant Journal **13,** 465-473.

Rudolf, M., Machettira, A.B., Gross, L.E., Weber, K.L., Bolte, K., Bionda, T., Sommer, M.S., Maier, U.G., Weber, A.P.M., Schleiff, E., and Tripp, J. (2013). *In vivo* function of Tic22, a protein import component of the intermembrane space of chloroplasts. Molecular Plant **6,** 817-829.

Ruggiano, A., Foresti, O., and Carvalho, P. (2014). Quality control: ER-associated degradation: protein quality control and beyond. Journal of Cell Biology **204,** 869-879.

Rytz, T.C., Miller, M.J., McLoughlin, F., Augustine, R.C., Marshall, R.S., Juan, Y.T., Charng, Y.Y., Scalf, M., Smith, L.M., and Vierstra, R.D. (2018). SUMOylome profiling reveals a diverse array of nuclear targets modified by the SUMO ligase SIZ1 during heat stress. Plant Cell **30,** 1077-1099.

Sadali, N.M., Sowden, R.G., Ling, Q.H., and Jarvis, R.P. (2019). Differentiation of chromoplasts and other plastids in plants. Plant Cell Reports **38,** 803-818.

Sambrook, J. (2001). Molecular Cloning : a Laboratory Manual. (Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press).

Saracco, S.A., Miller, M.J., Kurepa, J., and Vierstra, R.D. (2007). Genetic analysis of SUMOylation in *Arabidopsis*: conjugation of SUMO1 and SUMO2 to nuclear proteins is essential. Plant Physiology **145,** 119-134.

Schleiff, E., Soll, J., Kuchler, M., Kuhlbrandt, W., and Harrer, R. (2003). Characterization of the translocon of the outer envelope of chloroplasts. Journal of Cell Biology **160,** 541-551.

Schneeberger, K., Ossowski, S., Lanz, C., Juul, T., Petersen, A.H., Nielsen, K.L., Jorgensen, J.E., Weigel, D., and Andersen, S.U. (2009). SHOREmap: simultaneous mapping and mutation identification by deep sequencing. Nature Methods **6,** 550-561.

Schnell, D.J. (2019). The TOC GTPase receptors: regulators of the fidelity, specificity and substrate profiles of the general protein import machinery of chloroplasts. Protein Journal **38,** 343-350.

Schnell, D.J., Kessler, F., and Blobel, G. (1994). Isolation of components of the chloroplast protein import machinery. Science **266,** 1007-1012.

Schwartz, D.C., and Hochstrasser, M. (2003). A superfamily of protein tags: ubiquitin, SUMO and related modifiers. Trends in Biochemical Sciences **28,** 321-328.

Shanmugabalaji, V., Chahtane, H., Accossato, S., Rahire, M., Gouzerh, G., Lopez-Molina, L., and Kessler, F. (2018). Chloroplast biogenesis controlled by DELLA-TOC159 interaction in early plant development. Current Biology **28,** 2616-2623

Shi, L.X., and Theg, S.M. (2013). The chloroplast protein import system: from algae to trees. Biochimica et Biophysica Acta **1833,** 314-331.

Sohrt, K., and Soll, J. (2000). Toc64, a new component of the protein translocon of chloroplasts. Journal of Cell Biology **148,** 1213-1221.

Sriramachandran, A.M., and Dohmen, R.J. (2014). SUMO-targeted ubiquitin ligases. Biochimica et Biophysica Acta **1843,** 75-85.

Stahl, T., Glockmann, C., Soll, J., and Heins, L. (1999). Tic40, a new "old" subunit of the chloroplast protein import translocon. Journal of Biological Chemistry **274,** 37467-37472.

Stanga, J.P., Boonsirichai, K., Sedbrook, J.C., Otegui, M.S., and Masson, P.H. (2009). A role for the TOC complex in *Arabidopsis* root gravitropism. Plant Physiology **149,** 1896-1905.

Stengel, A., Benz, P., Balsera, M., Soll, J., and Bolter, B. (2008). TIC62 redox-regulated translocon composition and dynamics. Journal of Biological Chemistry **283,** 6656-6667.

Stevenson, J., Huang, E.Y., and Olzmann, J.A. (2016). Endoplasmic reticulum-associated degradation and lipid homeostasis. Annual Review of Nutrition **36,** 511-542.

Stewart, M.D., Ritterhoff, T., Klevit, R.E., and Brzovic, P.S. (2016). E2 enzymes: more than just middle men. Cell Research **26,** 423-440.

Sun, Y.J., Forouhar, F., Li Hm, H.M., Tu, S.L., Yeh, Y.H., Kao, S., Shr, H.L., Chou, C.C., Chen, C., and Hsiao, C.D. (2002). Crystal structure of pea Toc34, a novel GTPase of the chloroplast protein translocon. Nature Structural Biology **9,** 95-100.

Takahashi, Y., Iwase, M., Strunnikov, A.V., and Kikuchi, Y. (2008). Cytoplasmic sumoylation by PIAS-type Siz1-SUMO ligase. Cell Cycle **7,** 1738-1744.

Teng, Y.S., Su, Y.S., Chen, L.J., Lee, Y.J., Hwang, I., and Li, H.M. (2006). Tic21 is an essential translocon component for protein translocation across the chloroplast inner envelope membrane. Plant Cell **18,** 2247-2257.

Tomanov, K., Zeschmann, A., Hermkes, R., Eifler, K., Ziba, I., Grieco, M., Novatchkova, M., Hofmann, K., Hesse, H., and Bachmair, A. (2016). *Arabidopsis* PIAL1 and 2 promote SUMO chain formation as E4-type SUMO ligases and are involved in stress responses and sulfur metabolism. Plant Cell **28,** 1230-1230.

Tranel, P.J., Froehlich, J., Goyal, A., and Keegstra, K. (1995). A component of the chloroplastic protein import apparatus is targeted to the outer envelope membrane via a novel pathway. The EMBO Journal **14,** 2436-2446.

Tsai, J.Y., Chu, C.C., Yeh, Y.H., Chen, L.J., Li, H.M., and Hsiao, C.D. (2013). Structural characterizations of the chloroplast translocon protein Tic110. Plant Journal **75,** 847-857.

Tzfira, T., Tian, G.W., Lacroix, B., Vyas, S., Li, J., Leitner-Dagan, Y., Krichevsky, A., Taylor, T., Vainstein, A., and Citovsky, V. (2005). pSAT vectors: a modular series of plasmids for autofluorescent protein tagging and expression of multiple genes in plants. Plant Molecular Biology **57,** 503.

van den Burg, H.A., Kini, R.K., Schuurink, R.C., and Takken, F.L.W. (2010). *Arabidopsis* small ubiquitin-like modifier paralogs have distinct functions in development and defense. Plant Cell **22,** 1998-2016.

Vembar, S.S., and Brodsky, J.L. (2008). One step at a time: endoplasmic reticulum-associated degradation. Nature Reviews Molecular Cell Biology **9,** 944-957.

Vierstra, R.D. (2009). The ubiquitin-26S proteasome system at the nexus of plant biology. Nature Reviews Molecular Cell Biology **10,** 385-397.

Wallas, T.R., Smith, M.D., Sanchez-Nieto, S., and Schnell, D.J. (2003). The roles of Toc34 and Toc75 in targeting the Toc159 preprotein receptor to chloroplasts. Journal of Biological Chemistry **278,** 44289-44297.

Wang, Y., and Dasso, M. (2009). SUMOylation and deSUMOylation at a glance. Journal of Cell Science **122,** 4249-4252.

Watson, S.J., Sowden, R.G., and Jarvis, P. (2018). Abiotic stress-induced chloroplast proteome remodelling: a mechanistic overview. Journal of Experimental Botany **69,** 2773-2781.

Woodson, J.D., Joens, M.S., Sinson, A.B., Gilkerson, J., Salome, P.A., Weigel, D., Fitzpatrick, J.A., and Chory, J. (2015). Ubiquitin facilitates a quality-control pathway that removes damaged chloroplasts. Science **350,** 450-454.

Wu, C., Seibert, F.S., and Ko, K. (1994). Identification of chloroplast envelope proteins in close physical proximity to a partially translocated chimeric precursor protein. Journal of Biological Chemistry **269,** 32264-32271.

Wu, F.H., Shen, S.C., Lee, L.Y., Lee, S.H., Chan, M.T., and Lin, C.S. (2009). Tape-*Arabidopsis* Sandwich - a simpler *Arabidopsis* protoplast isolation method. Plant Methods **5,** 16.

Xiong, R.Y., and Wang, A.M. (2013). SCE1, the SUMO-conjugating enzyme in plants that interacts with NIb, the RNA-dependent RNA polymerase of turnip mosaic virus, is required for viral infection. Journal of Virology **87,** 4704-4715.

Xu, P.L., Yuan, D.K., Liu, M., Li, C.X., Liu, Y.Y., Zhang, S.C., Yao, N., and Yang, C.W. (2013). AtMMS21, an SMC5/6 complex subunit, is involved in stem cell niche maintenance and DNA damage responses in *Arabidopsis* roots. Plant Physiology **161,** 1755-1768.

Yoon, H.S., Hackett, J.D., Ciniglia, C., Pinto, G., and Bhattacharya, D. (2004). A molecular timeline for the origin of photosynthetic eukaryotes. Molecular Biology and Evolution **21,** 809-818.

Zhang, J.J., Lai, J.B., Wang, F.G., Yang, S.G., He, Z.P., Jiang, J.M., Li, Q.L., Wu, Q., Liu, Y.Y., Yu, M.Y., Du, J.J., Xie, Q., Wu, K.Q., and Yang, C.W. (2017). A SUMO ligase AtMMS21 regulates the stability of the chromatin remodeler BRAHMA in root development. Plant Physiology **173,** 1574-1582.

Zheng, N., and Shabek, N. (2017). Ubiquitin ligases: structure, function, and regulation. Annual Review of Biochemistry **86,** 129-157.

Zheng, Y., Schumaker, K.S., and Guo, Y. (2012). Sumoylation of transcription factor MYB30 by the small ubiquitin-like modifier E3 ligase SIZ1 mediates abscisic acid response in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the United States of America **109,** 12822-12827.

Zufferey, M., Montandon, C., Douet, V., Demarsy, E., Agne, B., Baginsky, S., and Kessler, F. (2017). The novel chloroplast outer membrane kinase KOC1 is a required component of the plastid protein import machinery. Journal of Biological Chemistry **292,** 6952-6964.