

**Elucidating the function of
the *suppressor of ppi1* locus 2**

William John Broad

Wolfson College

D.Phil. in Plant Sciences

Trinity 2017



Elucidating the function of the *suppressor of ppi1 locus 2*

William John Broad

Wolfson College

D.Phil. in Plant Sciences

Trinity 2017

Abstract

Plastids are a hub of biosynthetic activity, and in different tissues they take on specialised functions and develop into complex structures (e.g., chloroplasts in leaves). The structural and metabolic diversity and dynamism of plastids is dependent upon proteins imported from the cytosol. At the outer-envelope membrane of plastids, such proteins are transported by the translocon at the outer-envelope membrane of chloroplasts (TOC), comprising a protein-conducting channel and receptor proteins. Mutations affecting these TOC proteins cause profound effects to plastid development. To understand more about plastid development and the regulation of the TOC machinery, a screen for second-site suppressors of a mutant lacking the Toc33 receptor, called *plastid protein import 1 (ppi1)*, was previously conducted. In this thesis, the resulting *suppressor of ppi1 locus 2 (sp2)* mutants were characterised. The *SP2* genetic locus was identified, and it was found to affect the development of chloroplasts in a mechanism related to protein import. This was determined to involve a strong negative regulatory effect on the abundance of the TOC proteins, which was reminiscent of the effect of the *SP1* locus. *SP1* is an E3 ubiquitin ligase that triggers proteasomal degradation of TOC proteins. For this to occur, it is thought that TOC proteins require a mechanism of extraction from the chloroplast membrane, in a pathway that would require a retrotranslocon, as is the case for the degradation of endoplasmic reticulum (ER) proteins by the ER-associated degradation (ERAD) system. As a relative of the Omp85 superfamily of protein transporters, *SP2* is here proposed to fulfil the role of retrotranslocon in a hypothesised chloroplast-associated degradation system (ChlorAD). Genetic interaction data suggest that *SP1* and *SP2* act in a common pathway, while co-immunoprecipitation and in vitro pulldown results reveal a strong physical interaction between *SP1* and *SP2*, as well as an interaction between *SP2* and the ERAD-associated ATPase, CDC48. Overall, evidence presented here strongly supports the proposed ChlorAD mechanism.

Acknowledgements

I give my thanks to my supervisor, Professor Paul Jarvis, for all his advice, support, and productive conversation, which have aided me on the course of completing this D.Phil. I thank also the researchers, D.Phil. students and technicians in the Jarvis group, and all members of the Department of Plant Sciences, for all of their assistance, advice, and coffee-time entertainment. In particular, I would like to thank Dr Qihua Ling, who has provided me mentorship throughout the past four years, and who has never ceased to amaze us all, with his breadth of scientific knowledge and experimental expertise. I give many thanks too to the wonderful Fan Jia, who has supported me through the better part of this D.Phil., despite the long days and weekends spent in the laboratory, I could not have done this without her.

Statement of authorship

The following people made contributions to the work presented in this thesis:

- Professor Paul Jarvis:** Edited the review article which forms the basis of Chapter 1, General Introduction.
- Dr Qihua Ling:** Edited the review article which forms the basis of Chapter 1, General Introduction. Produced the construct p2GW7-HA-CDC48. He also contributed significantly to understanding the function of SP1, and to developing the ChlorAD hypothesis, most importantly the involvement of CDC48.
- Dr Amy Baldwin:** Conducted the original screen which identified *sp2* alleles *sp2-1* to *-3*, which are reported and characterised in this thesis.
- Natalie Allcock:** Performed transmission electron microscopy on plant tissues (Fig. 2.2)
- Dr Tijen Demiral:** Performed the in vitro chloroplast protein import assays (Fig. 2.3).
- Dr Raphael Trösch:** Prepared *sp2-1* to *-3* for whole genome sequencing.
- Dr Matts Töpel:** Conducted analyses of whole genome sequences, and provided candidate *sp2* loci (Table 2.1).
- Dr Binqun Huang:** Identified the allele *sp2-5*, as well as the locus of the *sp2-5* mutation (Fig. 2.6).

Abbreviations

AKR2	Ankyrin repeat protein 2
AP	Ammonium persulfate
ATI-PS bodies	ATG8-Interacting protein containing plastid-associated bodies
ATP	Adenosine triphosphate
BamA	Beta-barrel assembly machinery A
BLAST	Basic Local Alignment Search Tool
BSA	Bovine serum albumin
CCV	CV-containing vesicle
CDC48	Cell division control protein 48
cDNA	Complementary DNA
CDS	Coding DNA sequence
ChlorAD	Chloroplast associated degradation
CV	Chloroplast vesiculation protein
Da	Daltons
dCAPS	Derived cleaved amplified polymorphic sequence
Der1	Degradation in the ER protein 1
DMF	n,n-Dimethyl formaldehyde
DMSO	Dimethylsulfoxide
dNTP	Deoxyribonucleotide phosphates
DTT	Dithiothreitol
DUB	Deubiquitinase
E3	E3 ubiquitin ligase
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
eIF4E1	Eukaryotic translation initiation factor 4E locus 1
EMS	Ethyl methanesulfonate
ER	Endoplasmic reticulum
ERAD	ER-associated degradation
F _n	Filial generation
FC2	Ferrochelatase 2

GAPDH	Glyceraldehyde-6-phosphate dehydrogenase
gDNA	Genomic DNA
GFP	Green fluorescent protein
GST	Glutathione S-transferase
GTP	Guanosine triphosphate
HA	Human influenza hemagglutinin
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hrd1	Hmg-CoA reductase degradation ligase1
HRP	Horse-radish peroxidase
Hsp _n	Heat-shock protein of n kDa
IEM	Inner-envelope membrane
IMS	Intermembrane space
IP	Immunoprecipitation
IPTG	Isopropyl β -D-1-thiogalactopyranoside
Koc1	Kinase at the chloroplast outer-membrane 1
LB	Lysogeny broth
M _n	Mutagenized plant generation
MAD	Mitochondria-associated degradation
MAFFT	Multiple alignment using fast Fourier transform
MCMC	Markov chain Monte-Carlo
MCS	Multiple-cloning site
MES	2-(N-morpholino)ethanesulfonic acid
mRNA	Messenger RNA
MS	Murashige and Skoog
mSSU	Mature RUBISCO small-subunit
Npl4	Nuclear pore-localisation protein 4
OEM	Outer-envelope membrane
OEP	Outer-envelope protein
OMM	Outer-mitochondrial membrane
OMP	Outer-membrane protein
Omp85	Outer-membrane protein of 85 kDa
ORF	Open reading-frame

P36	Protein of 36 kDa
P39	Protein of 39 kDa (alternative name for SP2)
PCR	Polymerase chain-reaction
POTRA	Polypeptide transport-associated domain
ppi	Plastid protein import
pSSU	Pre-RUBISCO small-subunit
PUB4	Plant U-box protein 4
PVDF	Polyvinylidene fluoride
qPCR	Quantitative real-time PCR
RCB	RUBICO-containing body
RING	Really interestin new gene domain
ROS	Reactive Oxygen species
RT-PCR	Reverse-transcript PCR
RUBSICO	Ribulose-1,5-bisphosphate carboxylase/oxygenase
SAM	Sorting and assembly machinery
SAM50	SAM protein of 50 kDa
SAV	Senescence-associated vesicle
SD	Standard deviation
SDS	Sodium dodecylsulfate
SDS-PAGE	SDS-Polyacrylamide gel electrophoresis
SEM	Standard error of the mean
Slp1	Stomatin-like protein 1
SP1	Suppressor of ppi1 locus 1
SP1flex	SP1 with deleted TMDs
SP2	Suppressor of ppi1 locus 2
SPL	SP1-Like
SPP	Stromal processing peptidase
SUBA	SUBcellular localisation database for Arabidopsis proteins
T _n	Transformed plant generation
TAIR	The Arabidopsis information resource
TBE	Tris-buffered EDTA
TBS-T	Tris-buffered saline with Tween20

T-DNA	Transfer-DNA
TEAMP	Transient expression in Arabidopsis mesophyll protoplasts
TEM	Transmission electron microscopy
TEMED	Tetramethylethylenediamine
TGS	Tris glycine SDS buffer
TIC	Translocon at the IEM of the chloroplast
Tic _n	TIC protein of n kDa
TMD	Transmembrane domain
TOC	Translocon at the OEM of the chloroplast
Toc _n	TOC protein of n kDa
TOM	Translocon at the OMM
Tom40	TOM protein of 40 kDa
TPR	Tetratricopeptide repeat protein
TpsB	Two-partner secretion system B
Tris	Tris(hydroxymethyl)aminomethane
Uba	E1 ubiquitin activating protein
Ubc	E2 ubiquitin conjugating protein
Ubx	Ubiquitin regulatory X domain containing protein
Ufd1	Ubiquitin fusion degradation protein 1
UPS	Ubiquitin-proteasome system
WT	Wild type
w/v	$\frac{Weight\ (g)}{Volume\ (mL)}$
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Table of Contents

Abstract.....	i
Acknowledgements.....	ii
Statement of authorship.....	iii
Abbreviations.....	iv
General introduction	1
1.1 Introduction to plastids	3
1.1.1 Plastids: origin, morphology, and function	3
1.1.2 The plastid protein import system	5
1.1.3 Biogenesis of plastid outer-envelope membrane proteins.....	10
1.2 Regulation of plastids by the ubiquitin proteasome system	12
1.2.1 Introduction to the UPS.....	12
1.2.2 E3 ubiquitin ligases	14
1.2.3 General regulation of the plastid by the UPS.....	16
1.2.4 Actions of the UPS on the plastid protein import machinery	18
1.2.5 Actions of the UPS on plastids under abiotic stress	20
1.3 ER-associated degradation	25
1.4 Mitochondria-associated degradation	27
1.5 Aim of project	28
1.6 Figures and tables	29
Results I	33
2. Identification of <i>SP2</i>	35
2.1 Introduction	35
2.2 Results.....	41
2.2.1 Relation of <i>sp2</i> to chloroplast biogenesis	41
2.2.2 <i>SP2</i> is encoded by the locus At3g44160 in Arabidopsis	42
2.2.3 <i>SP2</i> is conserved in angiosperms.....	47
2.2.4 Analysis of <i>SP2</i> structure	48
2.2.5 The <i>SP2</i> -related gene, <i>P36</i> , is a pseudogene	50
2.3 Discussion	53
2.3.1 Evolution of <i>SP2</i>	53
2.3.2 <i>SP2</i> structure and function	55

2.4 Figures and tables	60
Results II.....	83
3. Functional analysis of SP2	85
3.1 Introduction	85
3.2 Results.....	87
3.2.1 SP2 specifically regulates the TOC components.....	87
3.2.2 Localisation and interactions of SP2	94
3.3 Discussion	101
3.4 Figures	105
Results III	123
4. Interactions between SP2 and SP1.....	125
4.1 Introduction	125
4.2 Results.....	129
4.2.1 Test of the similarity between <i>sp1</i> and <i>sp2</i> phenotypes	129
4.2.2 Analyses of genetic interactions between <i>SP1</i> and <i>SP2</i>	129
4.2.3 Analyses of the protein-protein interactions of SP2	133
4.3 Discussion	137
4.3.1 SP2 and SP1 function in a common pathway.....	137
4.3.2 The relationship of SP2 to the ChlorAD hypothesis.....	141
4.4 Figures	144
General discussion.....	155
5.1 The context of this work.....	156
5.2 Summary of key findings	159
5.3 Critical appraisal of the role of SP2 in ChlorAD.....	161
5.3.1 Assessment of the ability of SP2 to participate in ChlorAD	161
5.3.2 Alternative interpretations for SP2 function.....	165
5.4 Further perspectives	167
5.4.1 The influence of ChlorAD on specific plastid protein import pathways.....	167
5.4.2 Questions on the evolutionary origins of ChlorAD	169
5.5 Figures.....	171
Materials and methods.....	173
6.1 General use of Arabidopsis	174
6.1.1 Preparation of MS growth media	174

6.1.2 Germination of Arabidopsis seeds on MS media.....	174
6.1.3 Growth of Arabidopsis plants on soil.....	175
6.1.4 Salt-stress assay.....	176
6.1.5 Crossing of Arabidopsis plants.....	177
6.1.6 Floral-dip transformation of Arabidopsis plants.....	177
6.1.7 Chlorophyll measurements	178
6.1.8 Measurement of rosette diameter	180
6.1.9 TEM images	180
6.2 General use of bacteria	181
6.2.1 Preparation of LB medium.....	181
6.2.2 Heat-shock transformation of <i>E. coli</i> and plasmid isolation	181
6.2.3 Blue-white screening	182
6.2.4 Freeze-thaw transformation of <i>Agrobacterium</i>	182
6.3 DNA related techniques.....	184
6.3.1 PCR, agarose gel electrophoresis, and gel band purification ...	184
6.3.2 Restriction digestion	184
6.3.3 Cloning of a DNA sequence using the Gateway™ technology .	185
6.3.4 Cloning using digestion and ligation	187
6.3.5 Constructs made	188
6.3.6 Plant DNA extraction and genotyping.....	189
6.3.7 Verification of <i>sp2</i> point mutants	191
6.4 RNA-related techniques	192
6.4.1 RNA extraction and cDNA synthesis	192
6.4.2 Semi-quantitative RT-PCR	192
6.4.3 qPCR.....	193
6.5 Protein analysis.....	194
6.5.1 Total protein extraction from Arabidopsis plants	194
6.5.2 Protein quantification by Bradford assay	194
6.5.3 Preparation of SDS-PAGE gels	195
6.5.4 SDS-PAGE	196
6.5.5 Coomassie staining	196
6.5.6 Western blotting	196
6.5.7 Protein semi-quantification	198
6.6 Experiments using isolated chloroplasts	199

6.6.1 In vitro protein import into chloroplasts.....	199
6.6.2 Alkaline treatment of chloroplasts	200
6.6.3 Chloroplast protease protection assays	200
6.7 Methods for protein interaction experiments.....	201
6.7.1 Transient expression in Arabidopsis protoplasts	201
6.7.2 Co-immunoprecipitation	201
6.7.3 In vitro pulldown	203
6.8 Phylogenetic analysis.....	205
6.9 Tables.....	206
7. Bibliography	210

Chapter 1

General introduction

Statement of authorship

This chapter was adapted from Broad et al. (2016).

Professor Paul Jarvis: Contributed to the editing of the manuscript.

Dr Qihua Ling: Contributed to the editing of the manuscript.

The section on mitochondria, written by Dr Qihua Ling, was not included.

Summary

Plastids emerged as a result of an endosymbiotic relationship between a cyanobacterium-like organism and an early eukaryotic cell some 1.5 billion years ago. Today, plastids are an integral and defining feature of the plant cell. The functions of plastids in plants extend far beyond photosynthesis. They are at the nexus of many biosynthetic pathways, and act in environmental perception and interaction. In higher plants, which contain different tissue types, plastids take on different forms and functions in each tissue. In such cases, plastids are said to differentiate into different plastid classes or types, such as chloroplasts, amyloplasts or chromoplasts. This differentiation is not fixed, and like other organelles, plastids and their functions change as the cell does, and remain plastic, and as such there are many examples of plastid interconversions. In higher plants chloroplasts must develop from proplastids originating from the embryo which are propagated in the meristem. The form and function of plastids in a given tissue is dependent largely on the organellar proteome, and the action of the proteins which make up that proteome. Despite retaining a small extant genome, modern plastids must import about 3000 types of protein (not all of which are of cyanobacterial origin) post-translationally from the cytosol. Machinery, called the translocons at the outer/inner envelope of chloroplasts (TOC/TIC), have evolved to enable such protein import into plastids. Recently, through the discovery of the suppressor of *ppi1* locus 1 (SP1) E3 ligase, the ubiquitin-proteasome system (UPS) was found to regulate this import machinery, and thereby influence plastid development and homeostasis. The aim of this thesis is to expand on this knowledge on the regulation of plastid protein import, by the identification, characterisation and functional analysis of the SP2 locus.

1. General introduction

1.1 Introduction to plastids

1.1.1 Plastids: origin, morphology, and function

Plastids began their existence as a group of bacteria capable of oxygenic photosynthesis, and plastids of today appear to be related to extant cyanobacteria (Zimorski et al., 2014). This group of ancient cyanobacteria developed an endosymbiotic relationship with a eukaryotic cell approximately 1.5 billion years ago (Yoon et al., 2004), and the event gave rise to three known extant lineages, the glaucophytes, chloroplastida, and rhodophytes (Zimorski et al., 2014). It is the chloroplastida which gave rise to the Viridiplantae, which include all of the green-algae and land plants.

Throughout Viridiplantae, chloroplasts have the same general structure. They consist of three membranes, the outer-envelope membrane (OEM), the inner-envelope membrane (IEM), and the thylakoid membrane. These three membranes enclose three distinct compartments, the inter-membrane space (IMS), between the OEM and IEM, the stroma between IEM and thylakoid, and the thylakoid lumen enclosed by the thylakoid membrane.

Just as plant tissues are not always green, plastids do not always have the elaborate structures of the chloroplast, such as an intricately folded and ordered thylakoid membrane, neither are they always so large. Different plastid types develop in order to meet changing functional requirements as the cells and tissues of the plant differentiate. In land plants, all plastid types develop from small undifferentiated plastids, called proplastids, which have necessary

metabolic functions but only contain an undeveloped internal membrane system (Bräutigam and Weber, 2009). These proplastids originate from the embryo, and are most commonly maternally inherited in angiosperms, but can be paternally or biparentally inherited in some species (Matsushima et al., 2008). Once developed, plastids continue to sense and respond to their environment and may interconvert upon further cellular differentiation (Wise, 2006; Jarvis and López-Juez, 2013). For example, proplastids differentiate into etioplasts in cotyledons in the absence of light, and then into chloroplasts upon illumination, and finally into gerontoplasts when leaves senesce (Wise, 2006; Jarvis and López-Juez, 2013). And there are many other observed plastid interconversions (Jarvis and López-Juez, 2013). The most well studied plastid is the chloroplast, where photosynthesis takes place. Other important plant plastids include the chromoplasts, named for the carotenoid pigments that provide colour to fruits and flowers, and the amyloplasts which function primarily in starch synthesis and storage (Jarvis and López-Juez, 2013); see Table 1.1 (p. 29) for a summary of the different plastid types in plants.

Plant plastids are metabolically and biosynthetically diverse, contributing to the synthesis of amino acids, fatty acids, starch, chlorophylls, carotenoids, signalling molecules, and various precursors, as well as the assimilation of nitrogen and sulphur (Wise, 2006; Sakamoto et al., 2008). Another noteworthy function of plastids is in gravitropism, wherein specialized amyloplasts called statoliths in root and shoot apical meristems enable gravity sensing: The mechanism relies on the high density of these organelles, linked to the starch they contain, which causes them to sediment in the direction of gravity; this

information is interpreted by the plant, possibly through interaction of the statoliths with the actin cytoskeleton (Blancaflor, 2013; Sato et al., 2015).

The diversity of functions mediated by plastids means that they have particular importance in cellular homeostasis, and in the perception and responses to signals during development and stress. The latter point is especially noteworthy as photosynthesis is a major source of reactive oxygen species (ROS), which act in signalling but can easily become damaging if not regulated (Asada, 2006; Shapiguzov et al., 2012).

The biochemical functions and differentiation state of plastids are defined by the organellar proteome, which is itself influenced by the combined effects of gene expression, protein targeting and import, and anterograde and retrograde signalling.

1.1.2 The plastid protein import system

It is estimated that up to 3000 plastid proteins (~98% of the organellar proteome) are synthesized in the cytosol, necessitating mechanisms for the targeting and import of such proteins (Jarvis, 2008; Li and Chiu, 2010). Protein translocation across the outer and inner envelope membranes is mediated by the TOC and TIC (translocon at the outer/inner membrane of chloroplasts) complexes, respectively, in what has been called the 'canonical' protein import pathway (Shi and Theg, 2013b). Other 'non-canonical' pathways to the plastid, including a co-translational pathway through the endomembrane system have been described (Villarejo et al., 2005), but these are not well characterized (Jarvis and López-Juez, 2013). Proteins which enter the plastid via the canonical pathway are synthesized as preproteins carrying an N-terminal transit peptide,

which is cleaved upon entering the plastid stroma by the stromal processing peptidase (SPP; Richter and Lamppa, 2003).

The TOC and TIC complexes comprise multiple proteins, the first of which were identified in the 1990s in pea. Now, there are many known TOC and TIC components, and these are conserved across many plant and algal species (Jarvis and López-Juez, 2013; Shi and Theg, 2013b). Only the TOC components will be described here, as only these are of real relevance to this thesis.

The core components of the TOC apparatus are Toc34, Toc75 and Toc159 – the numbers indicate their respective molecular weights in kDa (Shi and Theg, 2013b). The stoichiometry of these components in the TOC complex, in the order Toc34:Toc75:Toc159, has been found to be 4-5:4:1 (Schleiff et al., 2003) or 3:3:1 (Kikuchi et al., 2006). Toc75 forms the protein-conducting channel of the complex. It is an Omp85-related protein (Outer-membrane protein of 85 kDa is a founding member of a superfamily of β -barrel outer-membrane proteins) comprising a conserved β -barrel domain at the C-terminal end, which forms the membrane integral portion of the protein, and an N-terminal polypeptide transport associated (POTRA) domain (Day et al., 2014). The POTRA domains face the intermembrane space, paralleling the situation in mitochondria and bacteria where the POTRA domains face the intermembrane space and periplasmic space, respectively (Paila et al., 2016; O'Neil et al., 2017).

Toc34 and Toc159 act as preprotein receptors (Richardson et al., 2014). They interact specifically with preproteins as they arrive at the chloroplast surface, and regulate the initial stages of translocation through Toc75 (Wang et al., 2008). Toc34 and Toc159 are each anchored in the plastid OEM by a C-terminal transmembrane anchor (in Toc159 this domain is unusually large and of

unknown configuration, and is termed the M-domain). Both proteins also possess a GTPase domain that protrudes into the cytosol and is often called the G-domain, while Toc159 additionally has an acidic domain at its N-terminus called the A-domain (Shi and Theg, 2013b).

Each TOC component typically has multiple isoforms in plants. In *Arabidopsis*, the relevant families are as follows; isoforms of Toc34: atToc33 and atToc34 (Jarvis et al., 1998; Gutensohn et al., 2000); isoforms of Toc159: atToc159, atToc132, atToc120 and atToc90 (Bauer et al., 2000; Ivanova et al., 2004; Infanger et al., 2011); and isoforms of Toc75: the major isoform, atToc75-III; and the minor isoform, atToc75-IV (Baldwin et al., 2005; Huang et al., 2011). Note that the 'at' prefix in each case denotes species of origin (in this case *Arabidopsis thaliana*), while the Roman numeral suffix indicates the chromosome on which the isoform is encoded. In successive chapters, unless otherwise stated, only *Arabidopsis thaliana* TOC genes/proteins will be discussed and so the 'at' prefix will typically be omitted. Additionally, when the roman numeral is dropped from Toc75, it will be assumed that Toc75-III, or the Toc75 family and homologues in general are discussed.

These different TOC isoforms allow for different TOC arrangements to import distinct groups of proteins, perhaps at different developmental stages or plastid types (Jarvis and López-Juez, 2013; Li and Teng, 2013; Demarsy et al., 2014). For example, atToc33, atToc159 and atToc75-III form the main translocon in chloroplasts, transporting photosynthesis-related preproteins, often called photosynthetic proteins (Fig. 1.1; Jarvis and López-Juez, 2013; Demarsy et al., 2014). By contrast, another translocon containing atToc34, atToc132/120 and atToc75-III is thought to be the main complex in non-photosynthetic plastids,

and to conduct other preprotein types such as those required for general organelle functions, metabolic pathways aside from photosynthesis, and maintenance (Fig. 1.1; Jarvis and López-Juez, 2013; Demarsy et al., 2014).

A wide variety of observations support the existence of multiple TOC configurations with different client specificities. Different TOC receptor isoforms have been observed to preferentially associate with each other in distinct groups, and can be pulled down together in those groups (Ivanova et al., 2004). Furthermore, different TOC receptor isoforms have been shown to have differing affinities for different preprotein clients (Gutensohn et al., 2000; Jelic et al., 2003; Ivanova et al., 2004; Dutta et al., 2014). Knockout mutant plants lacking the different receptors also have phenotypes that are consistent with the proposed specificities of the receptors, as summarized below (Hiltbrunner et al., 2001; Kubis et al., 2003; Constan et al., 2004; Kubis et al., 2004). It is also noteworthy that some preproteins and transit peptides favour certain plastid types, or plastids of different ages, whereas others are imported independent of plastid type or age (Yan et al., 2006; Teng et al., 2012).

The mutants of TOC components differ in the extent to which they affect the import of different groups of preproteins, or the biogenesis of different plastid types. Mutants of atToc33 (termed *plastid protein import1*, or *ppi1*) and atToc159 (*ppi2*) have pale-yellow and albino phenotypes, respectively, reflecting reduced import of photosynthesis-related proteins and impaired chloroplast biogenesis (Jarvis et al., 1998; Bauer et al., 2000; Kubis et al., 2003). The main Arabidopsis Toc75 isoform, atToc75-III, is expressed at a high level in all tissues, and mutations affecting this component are embryo lethal, indicating its central role in protein import and showing that plastids are critical to embryo development

(Baldwin et al., 2005). The role of atToc75-III has been studied in other tissues by use of a hypomorphic allele and inducible knockdown plants (Huang et al., 2011). Knockdown of atToc75-III caused abnormal development of chloroplasts and non-green plastids, and reduced import of both photosynthesis-related and non-photosynthesis-related preproteins. This indicated that atToc75-III is important, non-specifically for a broad range of plastid functions, and that it performs a general role in protein import. Depletion of atToc75-III also caused reduced abundance of atToc33 and atToc159, consistent with previous reports that Toc75 aids in the biogenesis of TOC components (Wallas et al., 2003; Huang et al., 2011).

Mutations affecting atToc132 or atToc120 do not cause strong phenotypes individually, but corresponding double mutants display severe chlorotic phenotypes owing to plastid biogenesis defects, which indicates functional redundancy between these two proteins (Ivanova et al., 2004; Kubis et al., 2004). Mutants of atToc34 (*ppi3*) have less severe phenotypes than *ppi1*, with the only obvious defects occurring in the roots, presumably due to defective development of non-photosynthetic root plastids (Constan et al., 2004). However, *atTOC34* expression can almost fully complement *ppi1*, revealing a degree of redundancy, while the double homozygous mutant genotype affecting both atToc33 and atToc34 causes embryo lethality (Jarvis et al., 1998; Gutensohn et al., 2000).

Thus, multiple TOC configurations may give specificity to protein import such that in different plastid types, at different stages of development, or under different conditions, the necessary proteins may be imported into the plastid in the correct abundance. If there were just a single type of TOC complex, competition effects might arise; for example, at the onset of photomorphogenesis,

highly-abundant photosynthesis-related proteins could conceivably out-compete other proteins (e.g., those required in lower abundance for more general organelle functions and maintenance) for entry into the plastid (Jarvis and López-Juez, 2013).

1.1.3 Biogenesis of plastid outer-envelope membrane proteins

All known chloroplast OEM proteins are encoded by nuclear genes (Lee et al., 2014). Like the plastid preproteins, these are translated in the cytosol, and then targeted to the chloroplast OEM for assembly, though the mechanisms for this are not well understood. Aside from Toc75, no known chloroplast OEM proteins have cleavable transit peptides typical of chloroplast preproteins, and are synthesised at their mature size (Oh et al., 2015). The targeting information must be integral to the amino-acid sequence, and/or emergent secondary and tertiary structures (Lee et al., 2014; Oh et al., 2015).

Chloroplast OEM proteins fall into two broad categories, those with α -helical transmembrane domains (TMD), and those which form a β -barrel structure from multiple β -strands. Chloroplast OEM proteins have been observed to associate with cytosolic chaperones, such as Ankyrin repeat-containing 2 protein (AKR2) and Heat-shock protein 70 (Hsp70), which can associate with the chloroplast (Dhanoa et al., 2010; Lee et al., 2014).

How TMD proteins are assembled into the membrane is not yet known, however there is evidence that the TOC may be required for targeting and insertion of some of these proteins (Wallas et al., 2003; Tu et al., 2004; Huang et al., 2011). β -barrel OEM proteins are likely first imported and then assembled. This is well established for Toc75 (Oh et al., 2015), and evidence suggests the

same for the close relative of Toc75, OEP80 (outer-envelope protein of 80 kDa; Day et al., 2014). This could also be true for the other chloroplast β -barrel OEM proteins. Chloroplast β -barrel proteins, OEP24 and OEP37, expressed in yeast were assembled into the outer-mitochondrial membrane (OMM), dependent on both the translocon at the OMM (TOM), and the sorting and assembly machinery (SAM; Ulrich et al., 2012). This therefore indicates that the proteins were first imported by TOM, then assembled from the IMS by SAM. This would suggest that in the chloroplast, the process could be dependent upon the TOC for import, and OEP80 for assembly.

OEP80, Toc75-III, and SAM component, SAM50, are Omp85 proteins and relatives of bacterial outer-membrane, β -barrel assembly machinery A (BamA). BamA is responsible for assembly of β -barrel proteins in the outer-membranes of gram-negative bacteria (Noinaj et al., 2014). It is thought that OEP80 has retained this ancestral function, like SAM50 in the mitochondria, while Toc75 has neofunctionalized to become a protein transporter (Huang et al., 2011; Stroud et al., 2011; Töpel et al., 2012). The POTRA domains of Omp85 proteins are crucial co-factors in β -barrel protein assembly in bacteria and mitochondria (Stegmeier and Andersen, 2006; Stroud et al., 2011; Noinaj et al., 2013). The POTRA domains of OEP80 likely face the IMS, as is the case for Toc75, Sam50 in mitochondria, and similar to those of bacterial Omp85 proteins, which face the periplasmic space (Noinaj et al., 2013; Dastvan et al., 2016; Paila et al., 2016). Considering the function of related proteins, and the topology of OEP80, chloroplast β -barrel protein assembly therefore most likely occurs from the IMS, though this process and co-factors involved are yet to be studied.

1.2 Regulation of plastids by the ubiquitin proteasome system

1.2.1 Introduction to the UPS

The ubiquitin-proteasome system (UPS) can specifically target proteins in the cell for degradation. The process involves the modification of a target protein by the addition of ubiquitin to signal and mediate its degradation by the cytosolic 26S proteasome, a multi-subunit proteolytic complex. Ubiquitin modification (referred to here as ubiquitination) is a post-translational modification of proteins that involves the ligation of the 76-amino acid protein, ubiquitin, most commonly to a lysine residue of the substrate protein by an iso-peptide bond. Ubiquitin and the UPS are highly conserved and ubiquitous in eukaryotes, play central roles in the regulation and dynamics of the cellular proteome, and are critical to cellular homeostasis and responses to developmental and environmental signals (Vierstra, 2009; Kleiger and Mayor, 2014).

Ubiquitin is bound to its substrate protein through a cascade of three enzymes. An E1 ubiquitin activating enzyme binds and activates ubiquitin in an ATP-dependent process; then, ubiquitin is covalently attached to an E2 ubiquitin conjugating enzyme which delivers ubiquitin to substrate proteins; an E3 ubiquitin ligase acts to provide substrate specificity, allowing the E2 to covalently bind the C-terminal glycine of ubiquitin to the amino side chain of a lysine of the substrate protein (Vierstra, 2009; Kleiger and Mayor, 2014).

Targeting of a substrate protein to the proteasome requires a polyubiquitin chain to be generated on the substrate (Thrower et al., 2000). In this process, successive ubiquitin monomers are ligated to each other, most commonly via

lysine-48. It is unclear whether polymerization can be catalysed by the E3 ligase associated with the target protein. It may be that polyubiquitin chains are formed while ubiquitin is still attached to the E2, during the transfer of ubiquitin between E1 and the substrate; alternatively, other factors may be involved (Walsh and Sadanandom, 2014).

Polyubiquitinated proteins are rapidly directed to, and degraded by, the 26S proteasome. This makes identification of UPS targets a challenge, particularly if the target is of low abundance in vivo. At the proteasome, the ubiquitinated protein is unfolded and enters the proteolytic complex while the polyubiquitin chain is removed by deubiquitinases (DUBs) on the outside of the proteasome. DUBs generate ubiquitin monomers from polyubiquitin, which are released into the cytosol to be reused. Substrate proteins enter the 20S core proteolytic complex via the 19S regulatory complex. Once inside, the proteins are hydrolysed into short peptides by a variety of proteolytic sites inside the proteasome, which are able to hydrolyse peptide bonds between different residues. Peptides are then released into the cytosol (Kleiger and Mayor, 2014).

Proteasomal degradation is not the only fate of ubiquitinated proteins; a variety of different ubiquitin modifications trigger different fates. For example, in monoubiquitination and multiubiquitination, a single ubiquitin monomer is added to one or multiple lysine residues on the target protein, respectively (Walsh and Sadanandom, 2014). In addition, ubiquitin polymerization via each of ubiquitin's seven lysine residues has been reported (these are: K6, K11, K27, K29, K31, K48 and K63; Kim et al., 2013), which enables the possibility of multiple, different polyubiquitin types, as well as branched ubiquitin polymers (Walsh and Sadanandom, 2014). Other consequences of ubiquitin modification include

directing large protein complexes or cell compartments to the vacuole or lysosome for autophagy (Li et al., 2015), regulation of protein localization or activity (Guerra and Callis, 2012), and regulation or repair of chromatin (Holt et al., 2015).

1.2.2 E3 ubiquitin ligases

The E3 ubiquitin ligases (E3 ligases, or 'E3s') are the key determinants of specificity in the ubiquitin proteasome system, enabling the system to target individual proteins and protein complexes with high specificity or in a more general manner. They are single proteins or protein complexes which can associate with target proteins to mediate transfer of ubiquitin. Their activity may be modulated by post-translational modifications such as phosphorylation or ubiquitination, regulatory partner proteins, or adapter proteins that can broaden, restrict or change the specificity and activity of the enzyme (Guerra and Callis, 2012; Sadanandom et al., 2012). Many Really Interesting New Gene (RING) and U-box type E3 ligases (see below) are subject to autoubiquitination and so regulate their own abundance (de Bie and Ciechanover, 2011; Guerra and Callis, 2012). Reflecting their importance in determining specificity, as well as the breadth and diversity of UPS targets, the model plant *Arabidopsis thaliana* has ~1400 genes encoding E3 ligase components, but only 37 E2 genes and two E1 genes (Vierstra, 2009).

Although extraordinarily diverse, E3 ligases fit into three broad categories: RING, U-Box, and Homologous-to-E6AP-C-Terminus (HECT). The RING and U-Box domains are protein interaction modules which recruit the relevant E2 (Guerra and Callis, 2012). The RING domain comprises a zinc-finger structure, while the U-box is a homologous domain that does not require zinc for its

structure to assemble (Guerra and Callis, 2012). The RING group can be subdivided into single-subunit and multi-subunit categories: Single-subunit RING E3s can direct ubiquitination of their substrates alone, whereas multi-subunit E3s are protein complexes (e.g., SCF-type E3 ligases) that comprise additional adapter proteins such as F-Box proteins for substrate interaction (Vierstra, 2009; de Bie and Ciechanover, 2011; Guerra and Callis, 2012; Jia et al., 2015). The HECT E3 ligases are monomeric, and unlike other E3s they covalently bind ubiquitin prior to its ligation to the substrate (Vierstra, 2009; de Bie and Ciechanover, 2011; Guerra and Callis, 2012).

E3 ligases can exist in the cytosol or the nucleus, or have been predicted or shown to have transmembrane domains that enable their localization to the endomembrane system, to the OMM, or to the plastid OEM (Deshaies and Joazeiro, 2009). Current evidence suggests that E3s and other component of the UPS do not exist internally within organelles, with the exception of the nucleus (von Mikecz et al., 2008). Therefore, regulation of endosymbiotic organelles by the UPS is mediated by outer membrane components or cytosolic interactors. In the case of the endoplasmic reticulum (ER), the ER-associated degradation (ERAD) system has evolved to retrieve misfolded proteins from the ER lumen, a process which is coupled to the UPS. This is similarly the case in mitochondria, where the UPS can extend as far as the intermembrane space (Bragoszewski et al., 2015; Wu et al., 2016). This has been recently characterised, and this phenomenon has now been dubbed mitochondrial associated degradation (MAD), by analogy to ERAD (Bragoszewski et al., 2015; Wu et al., 2016).

1.2.3 General regulation of the plastid by the UPS

The first evidence for a role of the UPS in controlling chloroplasts was the discovery of the interaction of a cytosolic E3 ubiquitin ligase, Arabidopsis C-terminus of HSC70-Interacting Protein (atCHIP), with plastid preproteins in the cytosol (Shen et al., 2007; Lee et al., 2009). Related to this, another study found that degradation of two plastid preproteins by the proteasome in vitro was dependent on the presence of their transit peptides (Sako et al., 2014). These observations reveal the presence of a quality control pathway for clearing away accumulated or aggregated plastid preproteins in the cytosol to prevent cytotoxic effects.

Proteomic studies aimed at characterising the Arabidopsis ‘ubiquitinome’ interestingly identified many ubiquitinated chloroplast proteins (Kim et al., 2013; Svozil et al., 2014). However, these results do not necessarily indicate direct regulation of plastids by the UPS as a transit peptide was found to be still attached for some of the proteins (implying that this modification occurred in the cytosol prior to import), although not for all them. Moreover, many of the detected proteins function inside the plastid, and so it is unlikely that they were ubiquitinated at their final, functional destination. Other identified proteins were atToc159, atToc33 and atToc34, and there is substantial evidence for the ubiquitination of these proteins *in situ* at the chloroplast envelope, as discussed later; these proteomic studies support the ubiquitination of these TOC proteins in vivo. Another study specifically attempted to define the chloroplast outer-membrane proteome and discovered E3 ligases (Simm et al., 2013). However, the significance of these results is uncertain as some identifications could potentially have been due to sample contamination.

Other studies do support the existence of E3 ligases in plastids. The first, albeit tentative case was the RING-type E3 ligase, Shoot Gravitropism9 (SGR9; Nakamura et al., 2011). SGR9 was shown to localize to the amyloplasts in root endodermal cells by analysis of a green fluorescent protein (GFP) fusion of a point mutant form, SGR9^{W244A}, with disabled ubiquitin ligase function. It was not possible to assess localization of the wild-type SGR9-GFP fusion, which was thought to be due to instability of the protein as RING-type E3s often auto-ubiquitinate to regulate their own abundance (Deshaies and Joazeiro, 2009; de Bie and Ciechanover, 2011); inactivation of the ubiquitination function stabilized SGR9 allowing fluorescence imaging. However, there is no further information on the in vivo sub-plastidic localization or topology of the SGR9 protein. There is some uncertainty concerning its role due to the fact that it is predicted to have a transit peptide (Ling and Jarvis, 2015a): OEM proteins do not typically possess a transit peptide, and an E3 ligase would be unlikely to function properly inside the plastid due to its isolation from other UPS components. That said, the biogenesis of plastid OEM proteins is not well understood; while most insert without the aid of a transit peptide, the Toc75 protein does possess a transit peptide, and it uses a stop-transfer mechanism for its assembly into the OEM (Inoue et al., 2001; Inoue and Keegstra, 2003; Day et al., 2014).

The *SGR9* gene was identified in a mutagenesis screen for gravitropism mutants: *sgr9* plants were found to have defects in amyloplast sedimentation, a phenomenon that is thought to enable gravity sensing by the cell. Nakamura et al. (2011) found that an actin filament mutant, *fiz1*, which induces fragmentation of actin, suppressed the gravitropism defect of *sgr9*. They hypothesized that SGR9 facilitates amyloplast detachment from actin filaments to allow their

sedimentation, although the targets of SGR9 are currently unknown.

Interestingly, another study found that atToc75 and atToc132 (which are thought to be ubiquitinated as a part of a different regulatory mechanism; see later) have important roles in gravitropism (Stanga et al., 2009). There is also evidence to indicate that the TOC machinery interacts with actin (Jouhet and Gray, 2009), and so it is tempting to speculate that translocon components are targets of SGR9. However, it is equally possible that SGR9 operates through a different mechanism entirely.

1.2.4 Actions of the UPS on the plastid protein import machinery

The first plastid E3 ligase to be identified with clearly established localization and topology was Suppressor of *ppil* locus1 (SP1; Ling et al., 2012). SP1 was discovered in a mutagenesis screen designed to identify novel regulators of plastid protein import; i.e., second-site suppressors of *ppil* (the atToc33 mutant) were sought. As mentioned earlier, *ppil* has reduced ability to import photosynthetic proteins and a distinctive pale-yellow phenotype, but it is still able to complete its life-cycle (Jarvis et al., 1998). The *sp1* mutation was found to partially recover *ppil* by allowing increased protein import, resulting in better developed chloroplasts and a greener plant. SP1 was predicted to have two transmembrane domains, and biochemical analysis showed it to be an integral component of the plastid outer membrane, with a C-terminal, cytosol-facing RING domain and an intermembrane space domain connecting the two transmembrane domains (Fig. 1.1). The cytosolic orientation of the RING domain gives SP1 access to other components of the UPS. Substitution of a critical residue in the RING domain interfered with ubiquitination of SP1 in vitro and in vivo, which led to the conclusion that SP1 is regulated by auto-ubiquitination like other

E3 ligases (de Bie and Ciechanover, 2011; Ling et al., 2012). The intermembrane space domain of SP1 was shown to be important for interaction with target proteins.

The identification and characterization of SP1 revealed for the first time that the UPS acts directly upon plastids to regulate their development, which it does by targeting components of the TOC machinery for proteasomal degradation (Fig. 1.1; Ling et al., 2012). Analysis of *sp1* mutant and *SP1* overexpressor Arabidopsis plants showed that *SP1* expression levels correlate inversely with the abundance of TOC proteins in vivo. Co-immunoprecipitation and in vitro pulldown results showed that SP1 can interact with many Arabidopsis TOC proteins (atToc159, atToc132, atToc120, atToc34, atToc33 and atToc75), and as noted earlier such interaction was mediated primarily by the SP1 intermembrane space domain.

The SP1 protein was shown to mediate ubiquitination of atToc159, atToc75 and atToc33 in vitro, while the extent to which these components were ubiquitinated in vivo was dependent on the expression level of SP1 (Ling et al., 2012). Evidence that ubiquitinated TOC components are directed to the proteasome for degradation was provided by the observation that polyubiquitinated atToc33 is stabilized upon addition of the proteasome inhibitor MG132. The link between protein import and the proteasome was further supported by the observation that the phenotypes of *ppi1* and an atToc75-III hypomorphic mutant allele, *toc75-III-3*, are suppressed by the proteasome mutations *pbe1* or *rpn8a*, essentially phenocopying the suppression mediated by *sp1* (Ling et al., 2012).

The data described above, when considered alongside hypotheses discussed earlier concerning the roles of different TOC component isoforms in client-specific import pathways, suggested that SP1 could be a part of a mechanism that regulates TOC component abundance during important developmental transitions (e.g., during etioplast-to-chloroplast conversion in de-etiolation, or chloroplast-to-gerontoplast conversion in leaf senescence): with the assistance of SP1, TOC complexes might rearrange to contain different component isoforms, allowing different proteins to be imported as required (Fig. 1.1). Evidence for this hypothesis was gained by carefully analysing de-etiolation and leaf senescence in *sp1* mutant and SP1 overexpressor plants: in *sp1* the rates of transition were slower than in wild type, whereas in the overexpressor plants the transitions happened more quickly (Ling et al., 2012). During de-etiolation, the abundance of TOC components was found to change rapidly in an SP1-dependent manner; most notably, the ratio of atToc159 to atToc132/120 increased markedly in the wild type, but hardly at all in the *sp1* mutant, possibly to enable the biogenesis of photosynthetic proteins which rapidly accumulate during de-etiolation (Ling et al., 2012).

1.2.5 Actions of the UPS on plastids under abiotic stress

Further work on the function of SP1, revealed an important role for plastid protein import regulation in plant responses to abiotic stress. Ling and Jarvis (2015b) observed that SP1 overexpressor plants display greater tolerance of osmotic, salinity and oxidative stresses, as determined by measuring seedling greening, development or survival. The observed stress tolerance was found to be negatively correlated with ROS production (Fig. 1.2A), which can over-accumulate in chloroplasts under conditions which cause imbalances in

photosynthetic electron transport, such as the aforementioned abiotic stress conditions. In this case, the abundance of the H₂O₂ (a particularly stable ROS) was negatively correlated with SP1 abundance under the stress conditions tested (Ling and Jarvis, 2015b). To further support this connection, *sp1* was crossed to two mutants related to a ROS scavenging pathway, *sos2* and *enh1*. The *sp1 sos2* and *sp1 enh1* double mutants were found to display synergistically enhanced sensitivity when subjected to mild oxidative stress, a result that was interpreted to indicate that SP1 co-acts in ROS regulation along with SOS2 and ENH1; as SOS2 and ENH1 are known to operate in ROS scavenging, SP1 was proposed to operate in ROS avoidance (Ling and Jarvis, 2015b).

The accumulation of ROS is due to the overproduction of, and/or inability to use, energized electrons within the photosystems. ROS quenching pathways exist in the chloroplast, but their capacity can be exceeded under stress conditions (Shapiguzov et al., 2012; Suzuki et al., 2012). Ling and Jarvis, (2015b) hypothesized that SP1 helps to control ROS accumulation by limiting the import of photosynthetic preproteins into chloroplasts, by targeting TOC components for removal under stress conditions, thereby limiting the maintenance and biogenesis of the photosystems and reducing photosynthetic activity (Fig. 1.2A). This notion was supported by the observation that TOC components are rapidly and specifically depleted in Arabidopsis plants during stress treatments, and that such changes are dependent on SP1: TOC components reached an even lower level in SP1 overexpressor plants following stress, whereas in *sp1* mutant plants such changes were very much reduced (Ling and Jarvis, 2015b). These effects on TOC component abundance were correlated with changes in protein import capacity, which further supported the hypothesis (Ling and Jarvis, 2015b).

A complementary pathway of chloroplast maintenance under oxidative stress involving ubiquitination was revealed by Woodson et al. (2015), who identified a cytosolic E3 ubiquitin ligase, Plant U-Box4 (PUB4), and proposed its involvement in the selective, vacuolar degradation of whole, ROS-damaged chloroplasts by an unknown mechanism (Fig. 1.2B) As mentioned earlier, mitochondrial autophagy is closely associated with ubiquitination, which provides a targeting signal for the autophagic machinery. Damaged chloroplasts, or chloroplasts in dark-treated, senescing tissue, are known to be subject to recycling by autophagy in the vacuole (Wada et al., 2009; Ishida et al., 2014). However, information on the autophagy of whole organelles is limited, and it is better known that components of chloroplasts undergo autophagy via Rubisco-containing bodies (RCBs) or other vesicles called ATI-PS bodies (named for their association with the autophagy-related protein ATG8-Interacting1 [ATI1], and with plastids; Izumi, Hidema, & Ishida, 2015; F. Li & Vierstra, 2012; Xie, Michaeli, Peled-Zehavi, & Galili, 2015). A third pathway of chloroplast vacuolar degradation is not associated with autophagy, and is triggered by both senescence and abiotic stress. This pathway is dependent upon the Chloroplast Vesiculation (CV) protein, and involves the formation of CV-Containing Vesicles (CCVs) which mediate the transfer of stromal proteins, as well as thylakoid membrane and thylakoid lumen proteins, to the vacuole (Xie et al., 2015). Senescence Associated Vacuoles (SAVs) are a fourth pathway of chloroplast degradation, which exhibit independent proteolytic activity and are also independent of autophagy (Xie et al., 2015).

The Arabidopsis *pub4* mutant was identified as a suppressor of the *ferrochelatase 2* (*fc2*) mutant, which affects the plastid-localized Ferrochelatase 2

enzyme, a component of the tetrapyrrole biosynthetic pathway that differentiates the heme branch from the chlorophyll branch and synthesizes heme from protoporphyrin IX (Woodson et al., 2015). Protoporphyrin IX accumulates in *fc2* and is a potent photosensitizing molecule, generating the ROS singlet oxygen (Shao et al., 2007). Under short-day conditions (i.e., 8 or fewer hours of light per 24 hour period), such production of singlet oxygen leads to chloroplast damage and subsequent chloroplast degradation, as observed by electron microscopy, resulting in a failure to green and, ultimately, death of *fc2* mutants within three days (Woodson et al., 2015). In the *pub4 fc2* double mutant, plants are partially recovered and chloroplast number is increased, and such changes were attributed to the disabling of the selective degradation of damaged chloroplasts. Evidence for the PUB4-dependent ubiquitination of chloroplasts under stress was presented, but the *pub4* mutant does not share characteristics typical of autophagy-related mutants, indicating that PUB4 may operate via an alternative, unknown mechanism, or be a part of a highly specialized autophagic pathway. Possible effects of PUB4 on the abundance of known targets of ubiquitination, or indeed of any plastid proteins, have not been assessed, nor is there any indication as to what its targets might be in chloroplasts; PUB4 has not yet been shown to associate with plastid OEM proteins. Nonetheless, Woodson et al. (2015) present an interesting evidence for a new mechanism of ubiquitin-dependent regulation of chloroplast dynamics.

The discovery that chloroplast OEM proteins are targets of the UPS raises interesting questions. As these proteins are assembled into a membrane, they must also require a mechanism of removal. Ubiquitinated protein substrates are removed for degradation from the membranes of other organelles, such as the ER

and the mitochondria. These are: ERAD, and MAD, described in brief in the following sections. These two systems are analogous, and in addition, the ubiquitin-dependent protein translocation activity has even been reported in secondary plastids, of cryptophytes, alveolates, stramenopiles, and haptophytes. ERAD components were co-opted (but uncoupled from proteasomal degradation), to fulfil a role in protein import across the outermost membranes of plastids in these species, and is called the symbiont-derived ERAD-like machinery (SELMA; Zimorski and Martin, 2014; Gould et al., 2015). The ubiquity of the ERAD-like mechanism raises the interesting prospect that this operates at the chloroplast too, and is under current investigation in the Jarvis group. By analogy to ERAD, this is being termed chloroplast-associated degradation (ChlorAD).

1.3 ER-associated degradation

As the 26S proteasome is not present inside the ER, a system termed the ER-associated degradation (ERAD) system evolved for the retrotranslocation of misfolded proteins and certain other target proteins from the ER to the cytosol where the proteasome resides (Vembar and Brodsky, 2008; Ruggiano et al., 2014). ERAD is subdivided into three pathways, depending on where the substrate is relative to the ER membrane: ERAD-L, ERAD-M and ERAD-C. ERAD-M targets ER membrane proteins with misfolded transmembrane domains, and ERAD-C targets ER membrane proteins with misfolded cytosolic facing domains (Ruggiano et al., 2014). ERAD-L is the pathway which retrotranslocates soluble ER lumen proteins and ER membrane proteins with misfolded luminal domains (Ruggiano et al., 2014). This process requires multiple steps, as summarized below, and multiple components, which here are named according to the components characterized in yeast.

While pathways for specific substrates have been identified, the general recognition process for misfolded protein substrates is still unclear, although the Hmg-CoA reductase degradation ligase1 (Hrd1) complex at the ER membrane likely plays a crucial role (Ruggiano et al., 2014). The Hrd1 complex is believed to contain the following ER membrane proteins: Hrd1, an E3 ligase with a cytosolic RING finger domain; Degradation in the ER protein1 (Der1); and Hrd3 which has a lumen-oriented Tetratricopeptide Repet (TPR)-containing domain (Ruggiano et al., 2014). Recent work has shown that Hrd3 plays a crucial role in the recognition of luminal substrates (Mehnert et al., 2015). Once the substrate protein is in association with the Hrd1 complex at the ER membrane, its retrotranslocation across the membrane can be initiated. The identity of the

retrotranslocation is also uncertain, although recent evidence from in vitro studies on Hrd1 reveal that it alone can transport substrates across membranes in a ubiquitin-dependent manner (Baldrige and Rapoport, 2016), while other candidates have been proposed to fulfil this function, including the Sec61 complex and Der1 (Carvalho et al., 2006; Bolte et al., 2011; Ruggiano et al., 2014). During retrotranslocation, the substrate is eventually exposed to the cytosol, and is then ubiquitinated through a typical enzymatic cascade involving the E1 Uba1, the E2s Ubc1 and Ubc7, and the aforementioned E3 ligase Hrd1 (Bolte et al., 2011). The ER membrane protein Ubx2 recruits the AAA+ ATPase, Cell division control protein 48 (CDC48) to the ER membrane along with its cofactors: heterodimers of Npl4 and Ufd1 (Richly et al., 2005; Wolf and Stolz, 2012). Npl4 and Ufd1 are believed to be involved in substrate recruitment by CDC48, which then provides the motive force to retrotranslocate the substrate into the cytosol where it can be degraded by the 26S proteasome. Although well established in its role, the precise mechanism of translocation by CDC48 is still unclear (Wolf and Stolz, 2012; Ruggiano et al., 2014).

1.4 Mitochondria-associated degradation

Although it has been long known that E3 ligases, such as March1 and Mulan, are resident in the OMM, the elucidation of the mechanism of OMM protein removal, and the discovery that mitochondrial IMS proteins are also subject to quality control by the UPS, were much more recent developments (Bragoszewski et al., 2013; Bragoszewski et al., 2015; Wu et al., 2016).

Interestingly, and in marked contrast with ERAD, it was discovered that the beta-barrel protein Tom40 could fulfil the role of retrotranslocon for misfolded IMS proteins (Bragoszewski et al., 2015). On the other hand, it was shown later that, as in ERAD, the activity of CDC48 and its co-factors, Npl4 and Ufd1, are specifically linked to the abundance of OMM proteins (Wu et al., 2016).

Furthermore, this link was attributed to Doa1 acting as substrate-recruitment factor, fulfilling an analogous role to the Ubx proteins in ERAD (Wu et al., 2016; Zhang and Ye, 2016). Combined, what follows from these observations is an emerging mitochondria-associated degradation (MAD) pathway with clear mechanistic similarities to ERAD, and notable similarities and differences in the components involved, and how certain functions are fulfilled (Zhang and Ye, 2016).

1.5 Aim of project

The aim of this project was to build on the body of work summarized above, to better understand plastids, their development and dynamics, and the regulation of these processes. This will be achieved by characterising the *suppressor of *ppi1* locus 2* (*sp2*), identified from a mutagenesis screen which aimed to identify novel regulators of plastid protein import.

The phenotype of this mutation is rare in its ability to cause the suppression of a chloroplast biogenesis defect, and so it is thought that the gene encoded by the locus affected in *sp2* plants will in some way influence chloroplast biogenesis. How this component may fit in with our current view of chloroplast biogenesis is an important question which needs to be addressed.

In Chapter 2, phenotypes of *sp2 ppi1* plants will be characterised, and the locus affected by *sp2*, will be identified and the homology and function of the predicted sequence will then be explored. In Chapter 3 further characterisations will be made, using a variety of established genetic and molecular techniques that will aim to understand in more detail, the effects of *SP2* and its encoded protein on chloroplast biogenesis and its relation to plastid protein import. In Chapter 4, these characterisations, together with an understanding of the genetic identity of *SP2*, its homology, and available published information, will ultimately inform hypothesis driven experiments aimed at elucidating the interactions and function of *SP2*.

1.6 Figures and tables

Plastid Type	Description
Proplastid	Small, undifferentiated plastids found in egg (or pollen in certain species), meristem and embryonic cells, which contribute important biosynthetic functions (Bräutigam and Weber, 2009). May be present in root nodules of nitrogen fixing plants, where they function in nitrogen assimilation (Reynolds and Blevins, 1986).
Etioplast	Chloroplast progenitor organelles that develop in cells of dark-grown tissue and differentiate into chloroplasts upon illumination (Plösch et al., 2011).
Chloroplast	Green, photosynthetic plastids that are ubiquitous in palisade and mesophyll cells. They sense light and CO ₂ , and stress signals, and can drive the opening and closure of stomata in guard cells (Jarvis and López-Juez, 2013).
Gerontoplast	Chloroplast-derived organelles that catabolise the resource-rich components of chloroplasts during senescence (Wise, 2006; Golczyk et al., 2014)
Chromoplast	Organelles specialized for the synthesis and accumulation of carotenoid pigments, which function as attractants in flowers and fruits, to aid pollination and dispersal, and affect flavour and quality of fruit (Barsan et al., 2012).
Amyloplast	Synthesize and store starch in root, tuber and seed cells, and can function in gravity sensing in the root and apical meristems (Pogson et al., 2015; Sato et al., 2015).
Elaioplast	Plastids that synthesize lipids, for example for the exine during pollen development (Quilichini et al., 2014).

Table 1.1. Summary of Plant Plastid Types.

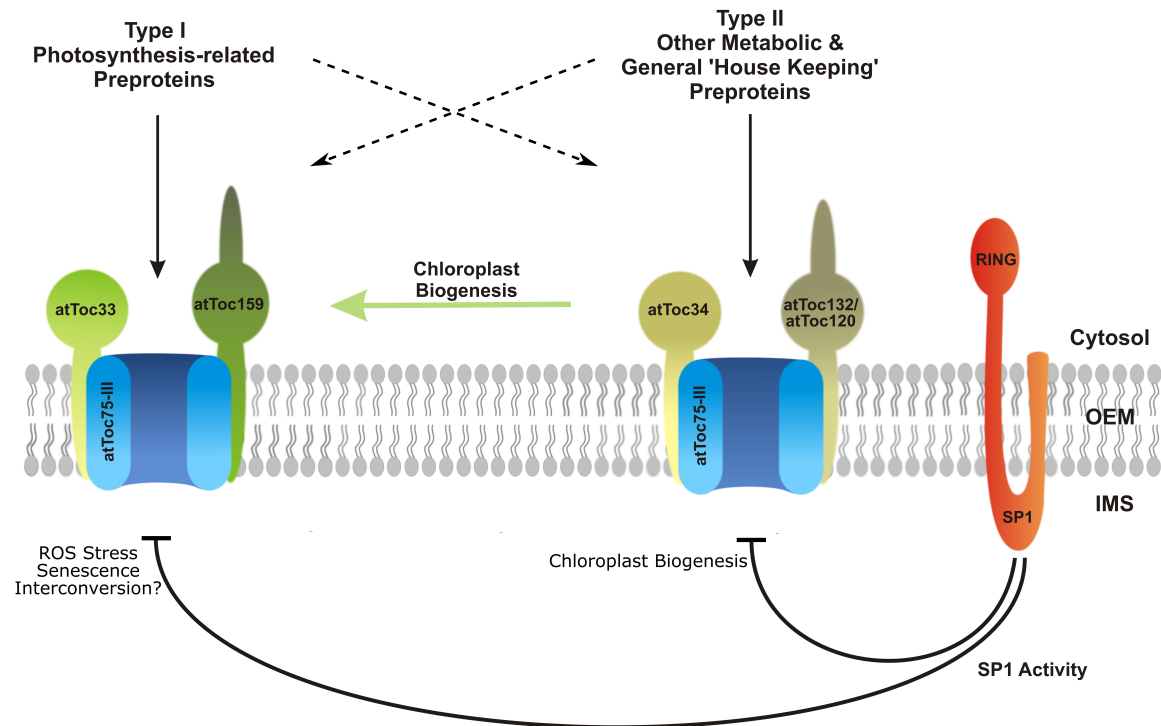


Figure 1.1. Regulation of protein import into plastids by the E3 ubiquitin ligase SP1. The figure shows a schematic representation of two outer-membrane translocon types: Type I consists of atToc33, atToc75-III, and atToc159, and is believed to preferentially transport highly abundant photosynthesis-related preproteins; Type II consists of atToc34, atToc75-III, and atToc132 and/or atToc120, and is believed to have preference for preproteins of other metabolic pathways and general functions of the plastid, such as plastid maintenance. At the onset of chloroplast biogenesis, the E3 ubiquitin ligase SP1 ubiquitinates components of the Type II TOC translocon to signal their proteasomal degradation, possibly to facilitate the assembly of components comprising the Type I translocon in the outer membrane. SP1 also ubiquitinates the Type I TOC complex during ROS stress, as part of a ROS avoidance strategy. Finally, SP1 may ubiquitinate TOC components to enable the transformation of chloroplasts into other plastid types, for example, during senescence, and plastid-type interconversions more generally. RING = really interesting new gene, an E3 domain which interacts with E2 ubiquitin conjugating enzymes; OEM = outer-envelope membrane of the plastid; IMS = intermembrane space. This was taken from (Broad et al., 2016)

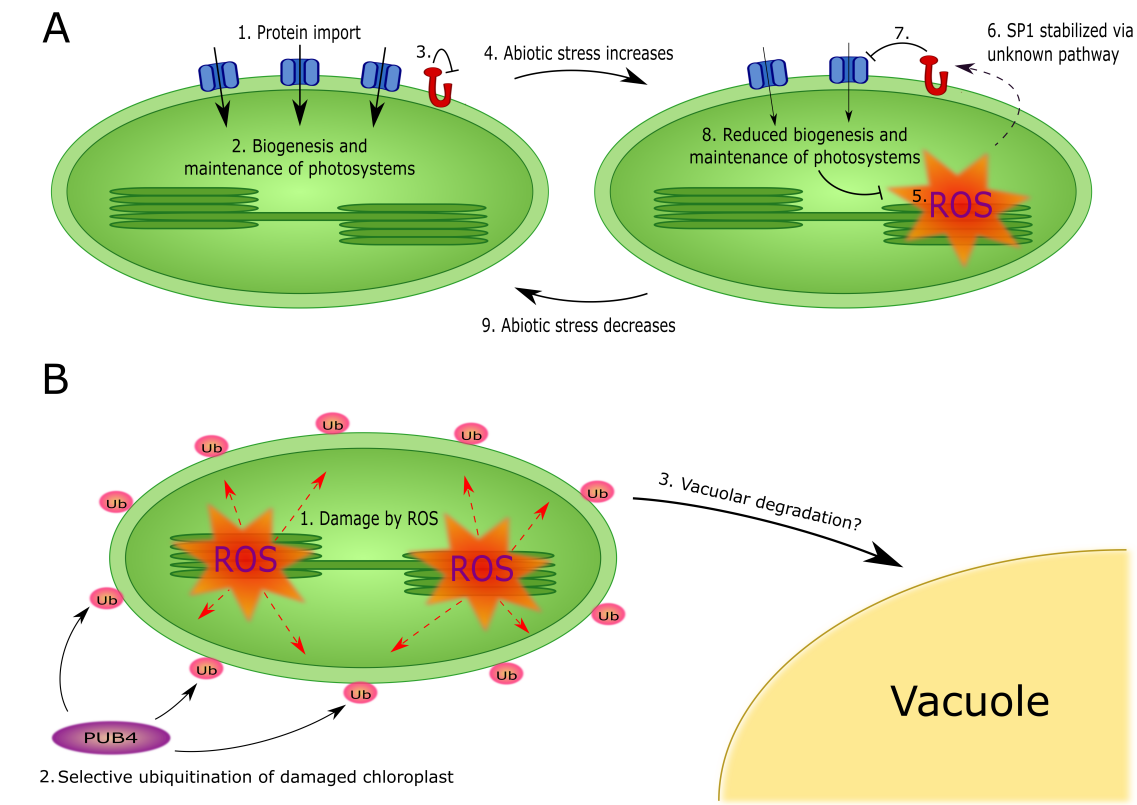


Figure 1.2. The regulation of chloroplasts under stress conditions by ubiquitin modification. (A) A mechanism for ROS avoidance is implemented under moderate abiotic stress (e.g., salinity or osmotic stresses). (1) Under normal conditions, proteins are imported into the chloroplast through the TOC translocon. (2) These imported proteins are required for the biogenesis and maintenance of the photosynthetic apparatus, as well as other metabolic functions. (3) Under such conditions, the OEM E3 ligase SP1 regulates its abundance by autoubiquitination. (4) Under moderate abiotic stress, chloroplast functions are affected. (5) Damaging ROS accumulate as a result of imbalances in photosynthesis. (6) In response, SP1 becomes stabilized via an unknown pathway. (7) SP1 is then able to target TOC components and so reduce protein import into the chloroplast. (8) Reduced protein import reduces the biogenesis and maintenance of the photosynthetic apparatus, and so photosynthesis is attenuated leading to reduced ROS accumulation. (9) When ROS accumulation is reduced, or the abiotic stress is alleviated, the system can return to normal. (B) A separate, quality-control mechanism selectively degrades damaged chloroplasts under more severe stress. (1) Some chloroplasts accumulate and are severely damaged by ROS under strongly photooxidizing conditions. (2) Through an unknown mechanism, certain plastid proteins (presumably in the outer-envelope

membrane) may become selectively ubiquitinated by the cytosolic E3 ligase PUB4. (3) This leads to the selective removal of these damaged chloroplasts through an unknown mechanism, possibly involving the vacuole. Ub = ubiquitin. This was taken from (Broad et al., 2016)

Chapter 2

Results I

Identification of *SP2*

Statement of authorship

Dr Amy Baldwin: Conducted the original screen which identified *sp2* alleles *sp2-1* to *-3*, which are studied in this thesis.

Natalie Allcock: Performed transmission electron microscopy on plant tissues (Fig. 2.2).

Dr Tijen Demiral: Performed the in vitro chloroplast protein import assays (Fig. 2.3).

Dr Raphael Trösch: Prepared *sp2-1* to *-3* for whole genome sequencing.

Dr Matts Töpel: Conducted analyses of whole genome sequences, and provided candidate *sp2* loci (Table 2.1).

Dr Binqun Huang: Identified the allele *sp2-5*, as well as the locus of the *sp2-5* mutation (Fig. 2.6).

Abstract

Plastids must import thousands of proteins from the surrounding cytosol. These are synthesised as preproteins, and at the plastid outer-envelope membrane they must pass through a machinery called the translocon at the outer-envelope of chloroplasts (TOC). There are different TOC types which are associated with the import of different sets of preproteins. The expression of different preproteins, and the requirements for different preproteins, changes in response to development, the type of tissue, and the environment experienced by the plant, and the TOC must be remodelled to reflect this. To search for novel regulators of the TOC apparatus, and of chloroplast biogenesis in general, the mutant of *Toc33*, *ppi1*, was subjected to a mutagenesis screen for second-site suppressors. Here, *suppressor of ppi1 locus 2* is described and characterised. Five alleles of *sp2* were described in this chapter, all of which suppress the *ppi1* phenotype, as measured by chlorophyll content. The suppression effect is related to chloroplast biogenesis, as chloroplast ultrastructure was recovered in *sp2 ppi1* plants compared to those of *ppi1*. This recovery of chloroplast biogenesis was linked to plastid protein import, as determined by an in vitro import assay. The *SP2* locus encodes an Omp85 superfamily β -barrel membrane protein which is conserved across angiosperms and closely related to OEP80, but is as yet of unknown function. SP2 lacks the POTRA domains typical of Omp85 proteins, but is still predicted to have the β -barrel structure. Consistent with this, critical amino acid motifs are conserved in SP2, and the *sp2* mutations *sp2-2* and *sp2-3*, are predicted to cause amino-acid substitutions which would disrupt the β -barrel structure given their location. Taken with other published observations, SP2 could have a transport function typical of Omp85 proteins.

2. Identification of *SP2*

2.1 Introduction

Plastids must import some 3000 proteins from the surrounding cytosol in order to develop and function (Jarvis, 2008; Sakamoto et al., 2008). Most proteins destined to the plastid are synthesised as preproteins with a cleavable transit peptide. Critical steps in the import of preproteins include their association with, and translocation across, the plastid membranes (Paila et al., 2015). At the outer-envelope membrane (OEM), these steps are facilitated by the TOC apparatus.

The plastid proteome requires remodelling according to plant and organelle development, the environment, and the metabolic needs of the cell (Jarvis and López-Juez, 2013; Paila et al., 2015). Evidence suggests that the TOC must be remodelled too, in order to facilitate efficient import of the required groups of preproteins (Ling et al., 2012; Demarsy et al., 2014; Ling and Jarvis, 2015b).

The TOC is composed of the protein conducting channel, Toc75-III, and the GTPase receptor proteins, Toc34 and Toc159, which were all originally discovered in pea chloroplasts (Shi and Theg, 2013b). Each of these are believed to have multiple isoforms in angiosperms, including Arabidopsis. These isoforms are associated with the import of specific subsets of chloroplast preproteins, and in some cases the expression of certain isoforms is related to the expression of genes encoding chloroplast preproteins (Yu and Li, 2001; Kubis et al., 2004; Kakizaki et al., 2009; Shi and Theg, 2013b). Certain TOC protein isoforms

preferentially associate, suggesting at least two types of translocon exist (Ivanova et al., 2004; Demarsy et al., 2014).

The translocon associated with the import of photosynthesis-related preproteins in *Arabidopsis* is composed of Toc33, Toc159 and Toc75-III, while a second type of translocon for the group of chloroplast preproteins required for general 'house-keeping' activities is thought to be composed of Toc34, Toc132 or Toc120, and Toc75-III (Kubis et al., 2003; Ivanova et al., 2004; Jarvis, 2008). The different TOC types are generally thought to dominate in certain plastid types, reflecting the profile of required preproteins, which is also evident in the mutant phenotypes of the individual TOC components (Jarvis et al., 1998; Bauer et al., 2000; Kubis et al., 2003; Kubis et al., 2004; Bräutigam and Weber, 2009; Stanga et al., 2009; Barsan et al., 2012; Demarsy et al., 2014). Moreover the expression of certain TOC proteins and their cognate preproteins may be intimately linked (Kakizaki et al., 2009). However these TOC complexes are not exclusive to one tissue, as the required preproteins are diverse, and so without multiple TOC types those preproteins which are highly abundant could potentially out-compete essential proteins of lower abundance for import into the plastid, and so impair plastid function (Jarvis and López-Juez, 2013; Ling and Jarvis, 2015a).

These pieces of evidence reveal that, as plastids interconvert between different types during organismal development, modulation of the TOC and its constituent proteins is necessary, and there are some more direct observations to support this. Teng et al. (2012) show that plastids of different ages have a differential ability to import preproteins with different transit peptides. This follows from early work conducted with plastids of different ages from wheat, which demonstrated a relationship between age and plastid-type on protein

import activity (Dahlin and Cline, 1991). Moreover, transit peptide dependent differential uptake of preproteins by chloroplasts and leucoplasts has also been demonstrated (Wan et al., 1996; Yan et al., 2006).

Furthermore, the energy economy of the cell may be a critical driving force behind the need to remodel the TOC. It has been estimated that an average of 650 ATP molecules are expended in the transport of a single preprotein through TOC and TIC (Shi and Theg, 2013a). This would mean that the saturating-rate of protein import would use approximately 0.6% of the maximum potential energy-output of the importing organelle (Shi and Theg, 2013a). Managing this expenditure could be especially critical during periods of stress, where the energy requirement of preprotein import and processing will amount to a greater proportion of energy output.

Chlorophyll and photosynthetic electron transport themselves can be a source of stress, for example by the production of reactive oxygen species (ROS). This is problematic particularly during times of abiotic stress, when the stomata close, causing a lack of reducing power for photosynthesis, due to a scarcity, or low concentrations of, CO₂. Therefore, at such times, import of photosynthesis-related preproteins could be reduced by remodelling and/or degrading the TOC thereby reducing import capacity, and reducing biogenesis and maintenance of the photosystems (Ling and Jarvis, 2015a; Broad et al., 2016), as proposed and evidenced by Ling and Jarvis (2015a). Additionally, there is evidence that temperature stresses could influence protein import into the chloroplast (Dutta et al., 2009). Moreover, components involved in quenching ROS, or otherwise responding to stress, could favour transport by a particular translocon, and so this could also drive remodelling of the TOC.

Recently it has become apparent that the chloroplast not only plays a role in biotic stresses, but may itself be a pathogenic target (de Torres Zabala et al., 2015). Pathogen effector molecules from bacteria and fungi have been localised to the chloroplast, with one even shown to have a cleavable transit peptide (Li et al., 2014; Petre et al., 2015; de Torres Zabala et al., 2015). As the TOC is implicated as a transporter of these molecules, it is tempting to speculate that it could itself be a pathogenic target, and that turn-over and remodelling of the TOC could play a role in biotic stress.

The above observations indicate that we should expect that the components of the TOC, and even whole TOC complexes, will be subject to regulation, be this at the gene-expression, transcriptional or post-translational level. However, currently there is little known about the regulation of the TOC components, and how protein import can be modulated.

To address these issues, a forward genetic screen was conceived to find novel regulators of chloroplast biogenesis. A mutant of chloroplast development was mutagenized to search for second-site suppressors of that mutation. Specifically, a mutant of Arabidopsis OEM TOC receptor, Toc33, was used. The mutant of Toc33, named *ppiL*, is distinctly pale-green in colour and stunted, yet remains fecund.

It was found that Toc33 is preferentially involved in the import of photosynthesis-related preproteins, though its homologue Toc34 is able to partially compensate (Jarvis et al., 1998; Kubis et al., 2003; Constan et al., 2004). This means that some degree of chloroplast biogenesis is possible in *ppiL*, which is indicated by the development of some thylakoid membranes (Jarvis et al., 1998; Kubis et al., 2003; Constan et al., 2004). This contrasts with the knockout

mutants of the other major receptor associated with import of photosynthesis-related preproteins (Toc159), called *ppi2*, which are almost completely albino, and completely lack thylakoids (Bauer et al., 2000; Kubis et al., 2004).

Because of its partially developed chloroplasts, and retained fecundity, the *ppi1* mutant was therefore an ideal candidate for a second-site suppressor screen to find new regulators of chloroplast biogenesis, specifically relating to the import of preproteins. The screen was conducted by the application of the mutagen ethyl methanesulfonate (EMS) to *ppi1* seeds. The resulting M2 were then screened for individuals with an increase in greening and/or size, as a surrogate phenotype for chloroplast biogenesis. Once suppressor mutants were identified, subsequent analyses, in particular electron micrographs of chloroplasts, were used to determine whether chloroplast development was recovered in these plants (i.e. whether the *ppi1* phenotype was suppressed).

The screen initially yielded four suppressor mutants, and these were discovered to be from two independent loci. One suppressor mutant was an allele of the *suppressor of ppi1 locus 1* (*sp1*; Ling et al., 2012), and three suppressor mutants were alleles of the *suppressor of ppi1 locus 2* (*sp2*; This study). In a subsequent, more extensive EMS screen, many more alleles of both *sp1* and *sp2* were identified (unpublished data), while in addition, several insertion alleles are available from the SALK T-DNA library (Ling et al., 2012; this study). In this thesis, only the first three identified *sp2* alleles, and the SALK T-DNA allele of *sp2*, are discussed in detail.

The *sp1* mutant has already been characterised in detail and described in the literature (Ling et al., 2012; Ling and Jarvis, 2015b). As discussed in Section 1.2.4, SP1 is an E3 ligase, a component of the ubiquitin-proteasome system

(UPS), which is a system of targeting proteins for degradation, via their post-translational modification with ubiquitin. SP1 was found to localise to the chloroplast OEM and to interact with TOC proteins in vitro and in vivo, and these interactions were determined to be mediated by the IMS domain of SP1 (Ling et al., 2012).

Thus, regulation of protein import into chloroplasts by SP1 occurs via the post-translational regulation of the TOC proteins. The biological relevance of this regulation has been demonstrated; it facilitates plastid type transitions, and helps to maintain plastid and photosynthetic homeostasis during abiotic stress (Ling et al., 2012; Ling and Jarvis, 2015b).

Prior to the initiation of this study, the genetic locus affected by the *sp2* mutation remained unidentified, and, because of this, only limited characterisations of the mutant had been made. Here, I present data concerning the genetic identity of *sp2*, its relationship to chloroplast biogenesis, and its homology and evolutionary relationships to other genes.

2.2 Results

2.2.1 Relation of *sp2* to chloroplast biogenesis

Three allelic variants of *sp2* were identified from the original *ppil* suppressor screen, and a fourth was subsequently obtained from the SALK T-DNA mutant library. The *sp2 ppil* plants are larger and greener than *ppil* (Fig. 2.1A), and these differences are apparent early in development (Fig. 2.1B). All of the *sp2* alleles showed a clear increase in chlorophyll content by fresh weight, when compared to the *ppil* single mutant (Fig. 2.1C). However, the different alleles are not equally able to suppress *ppil*; notably, in *sp2-2 ppil* and *sp2-3 ppil* the chlorophyll content increase was only 56%, and 66% respectively, whereas the *sp2-1* and T-DNA knockout allele *sp2-4* (Note: identification of the *sp2* locus and T-DNA allele will be described in Section 2.2.2) showed 100% and 104% increases in chlorophyll content, respectively, as compared to *ppil* (Fig. 2.1C).

There is a statistically significant difference in chlorophyll content between *sp2-1* and *sp2-2*, and *sp2-1* and *sp2-3* (Fig. 2.1C). This indicated the existence of ‘strong’ alleles that possibly knockout the gene or protein function, and hypomorphic alleles that provide a partial knockdown of the gene or protein function.

To test whether the suppression of *ppil* by *sp2* is related to chloroplast biogenesis, chloroplasts from *sp2 ppil* plants were analysed by TEM. Electron micrographs demonstrated an increase in size of the chloroplast, as well as a more highly developed ultra-structure (Fig. 2.2A). There was a clear increase in length, width, and cross-sectional area of the *sp2 ppil* chloroplasts compared to *ppil* (Fig. 2.4B, C, D). However, length divided by width remained unchanged

(Fig. 2.4D), indicating that *sp2* does not have an impact on the shape of chloroplasts.

Ultra-structure of the chloroplast is also more developed in *sp2 ppil* plants compared to *ppil*. Chloroplasts of *sp2 ppil* plants had an increased number of membranes in each granal stack, and an increased number of interconnections between granal stacks (Fig. 2.2F, G). These differences are significant, with $p = < 0.001$, for both measurements (Fig. 2.2F, G). However, development of the thylakoid membranes in *sp2 ppil* was still considerably less than in the wild type, as measured by the number of membranes in each granal stack, and the interconnections between stacks (Fig. 2.2F, G).

As *ppil* is a mutant of protein import into the chloroplast, a protein import assay was conducted to test whether the recovery of chloroplast biogenesis seen in *sp2 ppil*, mediated by *sp2*, is related to the rate of protein import into the chloroplast. The small subunit of the nucleus encoded ribulose-1,5-bisphosphate carboxylase/oxygenase (RUBSICO), with its native transit peptide, termed pre-RUBISCO small subunit (pSSU), was imported into isolated chloroplasts of *sp2 ppil*, and processed into mature-SSU (mSSU), at a higher rate than into those of *ppil* over a ten-minute time course (Fig. 2.3), although this was still slower than the import rate observed using chloroplasts of the wild type.

2.2.2 SP2 is encoded by the locus At3g44160 in Arabidopsis

2.2.2.1 Identification of the SP2 locus

The genetic locus of *sp2* was mapped to Arabidopsis chromosome number 3 via linkage mapping by Dr Raphael Trösch (data not shown). However, it was not possible to determine precisely the position of the genetic locus on

chromosome 3, and furthermore *sp2* appeared to be genetically linked with *ppi1-1*, which was unexpected as *ppi1* is located on chromosome 1. This was potentially due to an abnormality in the chromosomes of the *ppi1-1* allele chosen for the screen, which is a T-DNA insertion mutant (Jarvis et al., 1998). It has been reported that chromosomal rearrangements are common in T-DNA transformed plants (Clark and Krysan, 2010). To overcome these problems, a next-generation sequencing approach was taken to determine the identity of the genetic locus of *SP2*.

The three identified *sp2* alleles, *sp2-1*, *sp2-2*, *sp2-3*, and the genetic background *ppi1-1*, were subjected to whole genome shot-gun Illumina sequencing (The Genome Analysis Centre). The whole genome sequences were assembled by Dr Mats Töpel (University of Gothenburg). Loci having mutations within the open reading frame (ORF) in all three *sp2* alleles, but not the *ppi1-1* genetic background, were identified by Dr Töpel (Table 1).

The open reading frame At3g44160 (Berardini et al., 2015) was a strong candidate. The locus At3g44160 is predicted to encode a 362-amino acid long Omp85 protein, with a 16-strand β -barrel structure, which has high homology to chloroplast OEM protein OEP80. The predicted protein encoded by locus At3g44160 had already been identified in several phylogenetic studies as a close relative of OEP80 and the Toc75 family (Moslavac et al., 2005; Bredemeier et al., 2007; Töpel et al., 2012), and was recently described by Day et al. (2014) and Hsueh et al. (2017).

Each *sp2* allele contains a G-to-A mutation within the At3g44160 ORF (Table 2, Fig. 2.4A). These were likely to be the effects of EMS, which causes G-to-A and C-to-T substitutions (López-Juez and Hills, 2011). These mutations were

observed in all reads of the locus by shotgun sequencing (Table 2), as the sequenced DNA was from a pool of individual plants, this would rule out the possibility of allelic variation at the positions in question. Additionally, the number of reads at the relevant positions by whole genome sequencing was a minimum of 26 (Table 2), indicating that the candidate mutations were unlikely to have been the result of poor-quality sequences. These data were later substantiated by Sanger sequencing of PCR products amplified from the At3g44160 ORF, using newly extracted DNA from each mutant plant.

Each of the candidate mutations were predicted to have a consequence to the predicted encoded protein. The mutation in *sp2-1* is a G-to-A substitution in the sense strand of At3g44160, at position 2420 relative to the ATG, which is the final base of predicted intron 10 (Fig. 2.4A), and so it was predicted that this would cause a transcript splicing defect. This was demonstrated by RT-PCR, as a larger transcript was amplified as well as a transcript of approximately the expected size (Fig. 2.4B). The larger transcript was markedly less abundant than the transcript of expected size, and so these were respectively termed, minor and major forms of the transcript. Both of these transcripts were sequenced. This revealed that, in the minor transcript, a portion of predicted intron 10 is retained (Fig. 2.4C), leading to a missense mutation from position 310 of the protein, and truncation of the protein from position 328 (Fig. 2.5). Moreover, the major transcript contained a deletion of a G at position 920 in the CDS (Fig. 2.4C). This would correspond to the first G of exon 10, and its deletion causes a frame-shift and truncation of the predicted protein (Fig. 2.5).

The *sp2-2* and *sp2-3* mutations were both predicted to cause amino-acid substitutions. Allele *sp2-2* is a G-to-A substitution at position 2098 relative to the

ATG, and causes the amino-acid substitution R261H, while *sp2-3* is a G-to-A substitution at position 2281 relative to the ATG, and causes the amino-acid substitution G291R (Fig. 2.4A, 5). In both cases, there were no obvious impacts on transcript abundance in the mutants (Fig. 2.4B), indicating that the *sp2* phenotype in each case was caused by the amino-acid substitution, and protein function, rather than effects on the transcript.

In a second, larger *suppressor of ppil* mutagenesis screen, additional *suppressor of ppil locus 2* alleles have been discovered. One of these, named *sp2-5*, has been characterised (Fig. 2.6). This allele also has a G-to-A nucleotide substitution, occurring at position 34 relative to the ATG, at the start of intron 1, causing a splice-junction defect. Interestingly, in the transcript, this causes the deletion of six in-frame nucleotides (Fig. 2.6A), causing a deletion of two amino acids, G10 and R11, close to the N-terminus of the protein (Fig. 2.5). In the *ppil* background, *sp2-5* also has a significantly greater chlorophyll content than the amino-acid substitution mutant *sp2-2* (Fig. 2.6C), indicating the stronger effect of this mutation, and so could be a knock-out as is thought to be the case for *sp2-1* and *sp2-4*.

2.2.2.2 A T-DNA insertion in the locus At3g44160 can suppress *ppil*

To confirm that mutation of the ORF At3g44160 can indeed cause suppression of *ppil*, and to support the correct identification of At3g44160 as *SP2*, a T-DNA insertion mutant was identified from the SALK T-DNA library which contains an insertion in predicted exon 9 of At3g44160 (Fig. 2.4A, SALK_137135). In homozygous T-DNA mutant plants, the 5' portion of the *SP2* transcript was detected; however, attempts to PCR amplify the *SP2* transcript using primers flanking the putative insertion site yielded no detectable bands

(Fig. 2.4B). This indicated that the full-length *SP2* transcript was knocked out in these plants. The *sp2* T-DNA mutant, henceforth called *sp2-4*, was crossed with *ppil-2*. The choice of *ppil-2* for this cross was due to the apparent genetic linkage between the *sp2* alleles generated by EMS mutagenesis and the *TOC33* locus in *ppil-1*. In the resulting double homozygous *sp2-4 ppil-2* plants, the *ppil* phenotype was suppressed, as in the EMS-induced suppressor mutants (Fig. 2.1C). The suppression effect of *sp2-4* on *ppil* was most similar to that seen in *sp2-1 ppil*, as determined by chlorophyll content (Fig. 2.1C), which supports the assertion that *sp2-1* is a knockout allele while *sp2-2* and *sp2-3* are hypomorphic.

2.2.2.3 Overexpression of At3g44160 complements alleles of *sp2*

To further test the contribution of the mutation of At3g44160 to the phenotypic suppression of the *ppil*, and to begin to infer the function of this genetic locus, At3g44160 was overexpressed in *sp2 ppil* plants to test for phenotypic complementation. The mutants *sp2-1 ppil* and *sp2-2 ppil* were transformed with a construct containing the coding sequence (CDS) of At3g44160 placed downstream of the cauliflower mosaic virus 35S promoter, henceforth called *SP2OX*.

Semi-quantitative RT-PCR shows that there were increased levels of *SP2* transcript in the *SP2OX sp2 ppil* plants (Fig. 2.7C). The *SP2OX sp2-1 ppil* and *SP2OX sp2-2 ppil* plants appeared paler and smaller than untransformed *sp2-1 ppil* and *sp2-2 ppil* plants (Fig. 2.7A). These transformants were similar in appearance to *ppil*, and the chlorophyll contents of these plants had decreased to concentrations similar to those in *ppil* (Fig. 2.7B).

Overall, these results confirmed that the suppression of *ppi1* by *sp2* is a result of the lack of a functional copy of At3g44160.

2.2.3 *SP2* is conserved in angiosperms

Previous reports had indicated that *SP2* encodes an Outer-membrane protein 85 kDa / Two-partner-secretion B (Omp85/TpsB) superfamily protein, related to OEP80 and the Toc75 family (Moslavac et al., 2005; Bredemeier et al., 2007; Töpel et al., 2012; Day et al., 2014). However, the evolution and conservation of *SP2* in plants and extant cyanobacteria had not been studied specifically in these reports. To resolve these issues, BLAST searches for *SP2* homologues were conducted on species across the Viridiplantae, and phylogenetic trees were constructed to infer the evolution of *SP2* and related genes (Camacho et al., 2009).

BLAST searches of the Phytozome 12 collection for predicted proteins showed that Arabidopsis *SP2* has orthologues in most angiosperms, but is absent from the early diverging angiosperms sampled, *Spirodela polyrhiza* and *Amborella trichopoda* (Figs. 2.8, 9), and nor is it found in the gymnosperm, *Pinus taeda* (Day et al., 2014). The top BLAST hits for those, and other more distantly related species, return proteins which have greater sequence similarity to Arabidopsis OEP80, than to Arabidopsis *SP2*.

Phylogenetic analysis indicated a close relationship of *SP2* to OEP80, and that *SP2* may have arisen by a gene duplication occurring after the divergence of the sampled lycophyte, and likely early in the divergence of angiosperms (Figs. 2.8, 9). This divergence is consistent between both the Bayesian inference tree (Fig. 2.8), and the maximum likelihood tree (Fig. 2.9). Bayesian inference gives

strong support with a posterior probability of 0.99 for this node (Fig. 2.8). Maximum likelihood analysis gives a bootstrap of 0.79 (Fig. 2.9). This is not the most parsimonious result, as this would suppose that *SP2* was subsequently lost in some angiosperm species, including *A. trichopoda* and *S. polyrhiza*, though it is more parsimonious than a recent report which placed the divergence of OEP80 and SP2 between the divergence of land plants from algae (Day et al., 2014).

The branch subtending the SP2 clade is substantially longer than that subtending the OEP80 clade. The branch length would suggest that the Arabidopsis OEP80 sequence has undergone an average of 0.5 substitutions per site, whereas Arabidopsis SP2 sequence has undergone an average of 1 change per site, since their divergence (Fig. 2.8). This implies that after the duplication which gave rise to SP2, the SP2 sequence underwent an accelerated divergence, with respect to OEP80, and this may indicate neofunctionalization of SP2 (Figs. 2.8, 2.9; Töpel et al., 2012). Within the SP2 clade, as with the OEP80 clade, branch lengths are relatively short. This implies that between the represented species, SP2 is functionally conserved. However, it is worth noting that in cases where a species contained multiple predicted proteins for OEP80 or SP2, only the representative with highest homology to atSP2 was retained for phylogenetic analysis. Therefore, potentially more divergent homologues of atOEP80 and atSP2 are not represented.

2.2.4 Analysis of SP2 structure

2.2.4.1 SP2 is predicted to have a β -barrel structure

To further investigate the function of the SP2 protein, the consequences of the *sp2* mutations, and the relationship of SP2 to other Omp85 proteins, the SP2

sequence was analysed in relation to OEP80. The sequences of SP2 and OEP80 were aligned, and the structures of the SP2 and OEP80 proteins were predicted by homology modelling by Swiss-model (Biasini et al., 2014), using a crystal-structure of the related protein BamA (Noinaj et al., 2013; Day et al., 2014). The primary and secondary structures of SP2 are similar to those of OEP80, and the two proteins share 41% amino-acid sequence identity (Fig. 2.10). SP2 also retains two conserved motifs, called motifs 3 and 4, identified by Moslavac et al. (2005), in addition to the VRGY/F motif found in loop 6 of all Omp85 family members (Figs. 2.10, 11).

SP2 is different from most Omp85 proteins as it lacks POTRA domains at its N-terminal end, and contains no prominent domains apart from the β -barrel structure (Figs. 2.10, 11). Protein structure prediction by Swiss-model revealed that the first β -sheet of SP2 is predicted to begin at amino-acid 26, and that the protein has 16 predicted β -strands typical of Omp85 superfamily proteins (Figs. 2.10, 11). Therefore, apart from the β -barrel structure, SP2 only has a short leader sequence of 26 amino acids at its N-terminus. It is not the only Omp85 protein to lack POTRA domains, and in fact it is not the only chloroplast protein of this nature, as Toc75-IV similarly lacks POTRA domains (Baldwin et al., 2005; Heinz and Lithgow, 2014). However, details of the functions of POTRA-less β -barrel proteins remain elusive.

2.2.4.2 The effects of *SP2* mutants correlate to protein structure

To shed light on the importance of specific domains and residues of SP2, the positions of amino-acid substitutions caused by the *sp2* mutations were considered. The *sp2-2* mutation affects in the VRGY motif, mutating the R to H (Figs. 2.10, 11). This could be considered a relatively conservative substitution, as

both amino acids are basic at the pH's of the cytosol and plastid stroma (Su and Lai, 2017), suggesting that *sp2-2* is hypomorphic in comparison to the knockouts, *sp2-1* and *sp2-4*, although clearly it alters the function of a critical motif; the intermediate degree of suppression caused by this allele is consistent with this notion (Fig. 2.1). The *sp2-3* mutation lies in β -sheet 13, and results in a major change of the smallest amino acid, G, to a very large, positively charged residue, R (Figs. 2.10, 11). This substitution, therefore, may cause a critical structural fault in the β -barrel structure, resulting in reduced protein assembly or stability, and reduced function in *sp2-3*.

Interestingly, *sp2-5* causes the deletion of two amino acids (G10 and R11) preceding the first β -sheet, and yet causes a stronger suppression of *ppi1* than the substitution mutant, *sp2-2*, and similar suppression to the knockout allele *sp2-4* (Fig. 2.6). This indicates that these two amino acids, or the first 26 amino acids as a whole, are in some way critical to SP2 function.

2.2.5 The *SP2*-related gene, *P36*, is a pseudogene

Arabidopsis locus At3g48620 was previously named *AthP2* by Moslavac et al., (2005), and more recently *P36* by Nicolaisen et al. (2015). *P36* was originally identified in a phylogenetic study of plastid Omp85 proteins as being closely related to *SP2* and specific to *A. thaliana*, meaning it is not found in other land plants including the close relative, *A. lyrata* (Moslavac et al., 2005). Nicolaisen et al. (2015) also reported that *P36* evolved as a result of a recent duplication of *SP2*, and is specific to *A. thaliana*. However, the latter report additionally described mutant phenotypes of a *p36* T-DNA insertion line (termed *p36-1*), as well as *P36* gene expression patterns. The locus was studied alongside *SP2*, which was named in that paper as *P39* (Nicolaisen et al., 2015).

In contrast to the findings of Nicolaisen et al. (2015), I did not find any evidence of *P36* expression (Fig. 2.12). Firstly, I found that the primers used for quantitative real-time PCR (qPCR) by Nicolaisen et al. (2015), primer pair 1 in my study (Fig. 2.12), are not specific to *P36*; in total, over the two primers, only one position differs from the corresponding *SP2* sequence. As a result, they are able to efficiently amplify a product from the plasmid pDONR201 containing the *SP2* CDS, named in the figure as *SP2* plasmid (Fig. 2.12B). Moreover, these primers do not amplify a product of the size corresponding to the predicted *P36* transcript from wild type cDNA (Fig. 2.12B, Table 3): the single band observed is approximately 125 bp, the size predicted for *SP2*, whereas the band for *P36* is expected to be 212 bp, the same as the gDNA, as indicated by the arrowhead (Fig. 2.12B, Table 3). Thus, the reported *P36* qPCR data cannot be a true indication of *P36* gene expression.

In further analyses, it was not possible to specifically detect *P36* expression, or clone the putative *P36* CDS. Firstly, it should be noted that it is difficult to specifically amplify *P36* due to its very high homology with *SP2*. For example, primers designed to amplify full-length *P36* (primer pair 2; which over the two primers have only 2 mismatches with *SP2*) are also capable of amplifying *SP2* CDS from a plasmid, and the *SP2* gene from gDNA of the *p36-1* T-DNA line (Fig. 2.12B). This also happens when using a primer which should bind the 5' end of the *P36* CDS together with a primer in a central position beyond the *P36-1* T-DNA insertion site (Fig. 2.12B, primer pair 3); however, for these primers, the predicted amplicon sizes for the *SP2* and *P36* genomic sequences, and for the *SP2* and *P36* cDNA sequences, are very similar (Table 3), and so in this case it is not possible to rule out amplification of both amplicons.

Primers carefully designed to be highly specific for *P36*, containing very little homology to *SP2*, (the 8 most 3' nucleotides, of the forward primer are mismatched with respect to *SP2*) could amplify a band of the correct size from wild type gDNA, but not from wild type cDNA (Fig. 2.12B, primer pair 4; Table 3). These primers also could not amplify successfully from the *SP2* plasmid, and so one may conclude that the band amplified from the gDNA template is the genomic *P36* sequence. Thus, the failure of these primers to amplify from the wild-type cDNA was not due to suboptimal PCR conditions, and so the data strongly suggest that the *P36* gene is not expressed.

Primers specifically designed to amplify *P36* transcripts from wild-type cDNA template, by having them each bridging an intron, were unable to amplify products from any of the templates used, despite applying 40 PCR cycles (Fig. 2.12B, primer pair 5). This again strongly suggests that the *P36* gene is not expressed. However short of artificially generating the predicted *P36* CDS, there is no possible positive control available to test the efficiency of the primers used in this experiment.

Overall, I was not able to find any evidence of *P36* expression, and so I conclude that At3g48620/*P36* is a pseudogene. In this regard, it is noteworthy that the At3g48620/*P36* gene model differs significantly from that of its sister, *SP2*: specifically, it lacks three N-terminal exons, and additionally incorporates the sequence of an intron (corresponding to the final intron of *SP2*) into its largest final exon (Fig. 12A). These differences would be expected to impinge on the integrity of the protein's beta-barrel structure. The conclusion that At3g48620/*P36* is a pseudogene could be further supported in the future by conducting northern blot analysis using probes specific to *P36*.

2.3 Discussion

Given the diversity of TOC components present in land plants, including *Arabidopsis*, there is a distinct rationale to explore the regulation of these components, as well as chloroplast biogenesis and plastid protein import in general. Different TOC components have long been associated with the import of different types of preproteins, and different plastid types (Jarvis and López-Juez, 2013; Shi and Theg, 2013b; Demarsy et al., 2014; Ling and Jarvis, 2015; Ling and Jarvis, 2015b). Moreover, previously published observations concerning SP1 showed a clear connection of plant development and stress responses with the remodelling of the TOC complex (Ling et al., 2012; Ling and Jarvis, 2015b). Also identified in the screen which identified SP1 was SP2, which is described here.

As was the aim of the mutagenesis screen, the identified *sp2* alleles appear to have a relation to chloroplast biogenesis. This was determined by several observations; in addition to an increased chlorophyll content in comparison to *ppi1* (Fig. 2.1), *sp2 ppi1* plants also have recovered chloroplast ultrastructure (Fig. 2.2), and an increased capacity to import chloroplast preproteins (Fig. 2.3). The mechanism underpinning the relation of SP2 to chloroplast biogenesis and protein import into plastids will be the subject of successive chapters.

2.3.1 Evolution of *SP2*

It was discovered that *SP2* encodes an Omp85 protein, closely related to *OEP80*. *SP2* was shown to be conserved in most angiosperms, apart from the basal species, *Amborella trichopoda*, and *Spirodela polyrhiza* (Figs. 2.8, 2.9). Consistent with it only being found in the angiosperms, *SP2* appears to have evolved between the divergence of lycophytes and likely early in the divergence of

angiosperms (Figs. 2.8, 2.9). This is contrary to the findings of Day et al. (2014); their reported phylogenies suggest that *SP2* evolved in early land plants but only persisted in the direct ancestor to all but the most basal angiosperms, and was lost multiple times in all earlier diverging species sampled. This is not parsimonious, and it was conceded that it was contrary to their hypothesis. It was suggested the result arose due to confounding factors, such as a lack of sequence conservation in *SP2* (Day et al., 2014). The phylogenies presented in this report are therefore more consistent with the observed distribution of *SP2* homologues, though it must be recognised that only a small set of non-angiosperm sequences are available, and so it is possible that analysis of a more extensive set of species could yield a different result.

Although closely related, *SP2* is not simply OEP80 without POTRA domains, as there is considerable divergence between the amino-acid sequences of the two proteins: amino-acid sequence identity is 41%. This can also be seen clearly in the branch length subtending the *SP2* clade, which is considerably longer than that of OEP80, indicating neofunctionalization (Figs. 2.8, 2.9). *SP2* therefore seems to be functionally distinct from OEP80.

The short branch lengths within the *SP2* clade indicate a high level of conservation amongst *SP2* sequences (Figs. 2.8, 2.9). This further supports the notion of distinction from OEP80, and indicates broad functional relevance to the angiosperms. It further suggests that conclusions drawn from studies on *SP2* in *Arabidopsis* are also likely to have relevance to other important species.

The TAIR10 *Arabidopsis* genome contains a sequence closely related to *SP2*, which has been named, *P36*. This sequence is specific to *A. thaliana* and not present in other angiosperms. Nicolaisen et al. (2015) showed that this sequence

arose from a recent gene duplication. However, in addition to this, Nicolaisen et al. (2015) reported qPCR data revealing an expression pattern distinct from *SP2*. Furthermore, a T-DNA mutant containing an insertion in the *P36* coding sequence was characterised, and it was reported that, in this mutant, senescence is delayed, rosette area is reduced, chloroplast size in leaf and stem tissues is increased, and the number of starch granules in sepal plastids is decreased, all with respect to wild type (Nicolaisen et al., 2015).

The physiological analyses of this *P36* T-DNA mutant are of questionable relevance, as the data were representative of only a single genotype, which could have other confounding mutations. Moreover, further experimental work (such as complementation with the *P36* CDS) was not completed to show that mutation of *P36* was indeed responsible for the reported phenotypes. More importantly, in my studies on *P36*, I was unable to clone or amplify a cDNA specific to this locus (Fig. 2.12). In addition, I discovered that the qPCR primers used by Nicolaisen et al. (2015) to study *P36* expression were highly similar to *SP2*, differing by only one nucleotide, and the primer pair was only able to amplify a band of a size corresponding to *SP2*, and no band corresponding to *P36*. This calls into question the qPCR data of Nicolaisen et al. (2015), which clearly do not provide an accurate reflection of *P36* expression.

Based on the findings I present here, I conclude that the *P36* locus is not expressed to a detectable level, and is therefore most likely a pseudogene.

2.3.2 *SP2* structure and function

SP2 was discovered to be a member of the Omp85 superfamily of insertases and transporters found in the outer-membranes of gram-negative

bacteria, as well as the mitochondria and plastids of eukaryotes (Misra, 2012; Heinz and Lithgow, 2014). Like its relative, OEP80, SP2 is believed to be localised in chloroplasts. Day et al. (2014) were able to import radiolabelled, in vitro translated SP2 into isolated chloroplasts. And, using an antibody raised against Arabidopsis SP2, Hsueh et al. (2017) showed that SP2 immuno-reactive band co-migrates with Toc75 from Arabidopsis cellular subfractionation samples. Thus, it is expected that SP2 functions at the chloroplast OEM.

Omp85 proteins consist of a 16-strand, membrane-embedded β -barrel proximal to the C-terminus of the protein, and POTRA domains at the N-terminal end. The Omp85 protein in bacteria most closely related to those found in plastids is BamA. Proteins of this type have been studied extensively, and function in the assembly of β -barrel OMPs, by assembling unfolded proteins into the bacterial outer-membrane from the periplasm (Noinaj et al., 2013; Gu et al., 2016).

Toc75 and OEP80 are plastid Omp85 proteins. Toc75-III has neofunctionalized and has become a protein conducting channel, though it is still associated with the assembly of certain OEM proteins such as Toc159 (Wallas et al., 2003); in contrast, OEP80 likely retains the ancestral function similar to BamA (Hsu et al., 2008; Patel et al., 2008; Huang et al., 2011; Töpel et al., 2012; Paila et al., 2016). Both Toc75-III and OEP80 contain three POTRA domains (Töpel et al., 2012). The POTRA domains of OEP80 have not been studied, but the Toc75 POTRA domains are indispensable for the proper formation of the TOC complex, which contains Toc75, Toc33 and Toc159 with a stoichiometry of 3:3:1 or 4:4:1 (Ertel et al., 2005; Kikuchi et al., 2006). The Toc75 POTRA domains were

found to be required for interaction with other complex components, in addition to preproteins and chaperones like Tic22 (Paila et al., 2016; O'Neil et al., 2017).

In contrast with Toc75-III, in SP2 the POTRA domains are absent, as they are in Toc75-IV. Apart from the study which characterised Toc75-IV (Baldwin et al., 2005), and the two aforementioned reports mentioning SP2 (Day et al., 2014; Hsueh et al., 2017), there is currently very little information on 'POTRA-less' Omp85 proteins in the literature, despite many having been identified, almost exclusively in bacteria, in phylogenetic studies (Heinz and Lithgow, 2014). Therefore, there is very little information to indicate what the function of these proteins might be.

Arabidopsis mutants of Toc75-IV have some minor developmental defects particularly associated with etioplast-to-chloroplast transitions. However, the function of SP2 must be distinct from that of Toc75-IV, as a *toc75-IV* mutant was unable to suppress, or accentuate, the *ppi1* phenotype (Baldwin et al., 2005). Another example of a study on a POTRA-less Omp85 is work conducted with a POTRA deletion mutant of the mitochondrial protein, Sam50, which is responsible for assembling the mitochondrial outer-membrane protein import channel, Tom40. Sam50 only has one POTRA domain, but when this POTRA domain was deleted, Sam50 was still able to assemble Tom40, though in this case it was defective in substrate dissociation (Stroud et al., 2011). This indicates that POTRA domains may not be essential for insertase or transporter functions.

It is likely that SP2 forms a functioning protein with a mechanism similar to other Omp85 proteins. SP2 retains important motifs conserved across Omp85 proteins (Fig. 2.10). This includes the important VRGY sequence in loop 6, which, when mutated in *sp2-2*, evidently disrupts the protein's function (Figs. 2.1, 2.10,

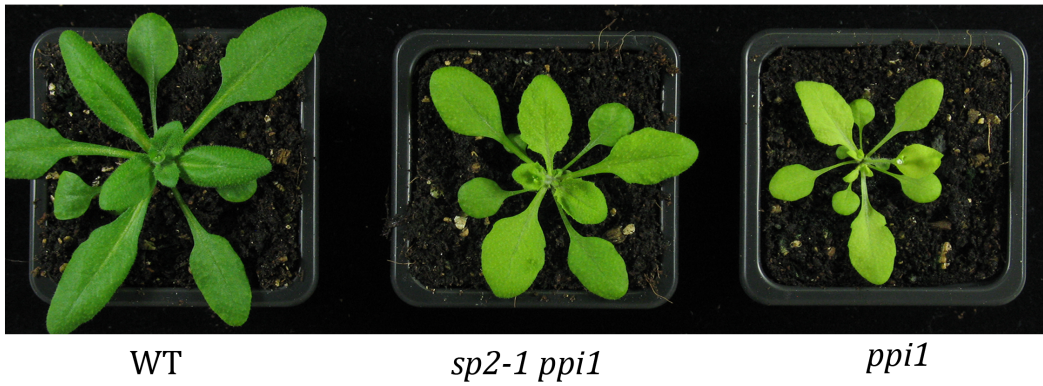
2.11). Structural studies tell us that loop 6 occludes the lumen of the β -barrel, and may function in substrate interaction, conformational changes, and/or gating activity of the channel by hydrogen-bonding between the VRGY motif on the loop with conserved residues from β -strands 12 to 15 on the luminal face of the β -barrel (Delattre et al., 2010; Noinaj et al., 2013; Maier et al., 2015; Hsueh et al., 2017). Hsueh et al. (2017) reconstituted SP2 into liposomes and tested the channel activity, and found that the size and properties of the channel were similar to those of OEP80. Moreover, the number of POTRA domains in Omp85 proteins is variable, and this may correlate with differences in the interaction partners or substrates of the protein. Therefore, SP2 may not be part of a canonical Omp85 complex, but could still retain a form of channel function related to the β -barrel structure, despite a lack of POTRA domains.

Although SP2 contains no POTRA domains, it does possess a 26-amino acid sequence preceding the first β -strand. The mutant *sp2-5*, which contains a two-amino acid deletion in this sequence, appears closer in phenotype to an *sp2* knockout allele than to an hypomorphic allele, indicating that in this mutant SP2 function is severely compromised (Fig. 2.6). Thus, interesting questions arise as to the role of this sequence, as it appears that it is essential for protein function. The role of this domain could be in targeting to the chloroplast. However, this seems unlikely as, other than Toc75-III which has a bipartite transit peptide (Endow et al., 2016), chloroplast β -barrel OEM proteins, like those from mitochondria, do not contain a transit peptide (Oh et al., 2015). Rather, the targeting information is thought to be intrinsic to the protein structure (Oh et al., 2015). Other possible reasons for the necessity of this domain relate to substrate

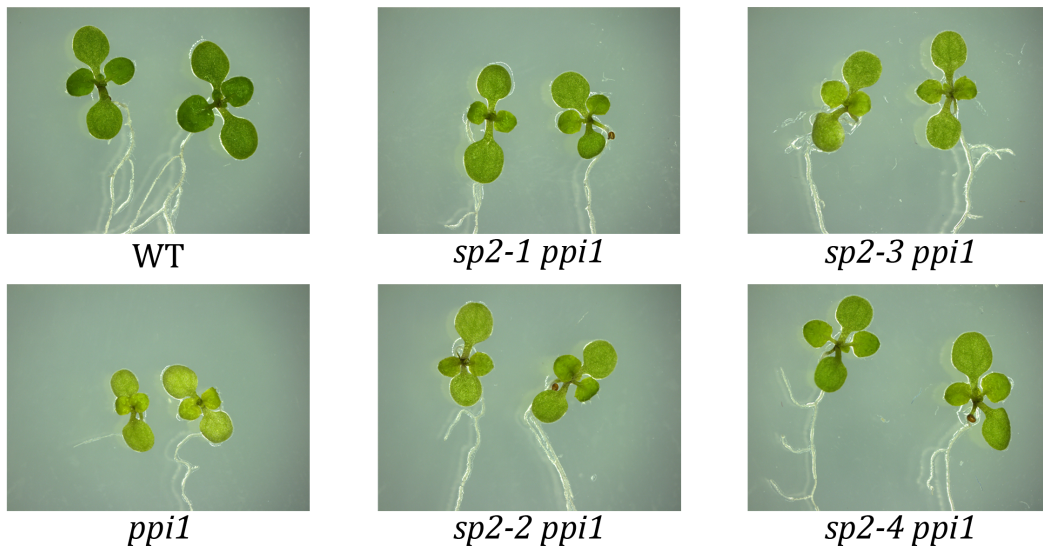
binding, interaction with assembly machinery, and gating activity of the channel.
A role in channel gating has already been proposed (Hsueh et al., 2017).

2.4 Figures and tables

A



B



C

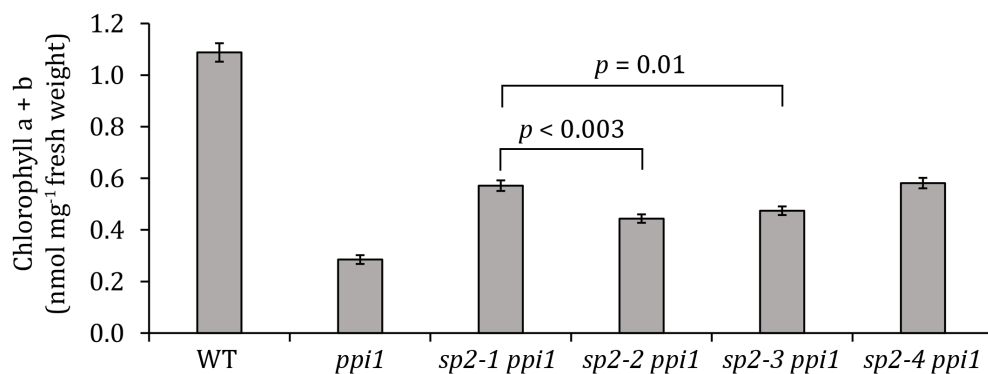


Figure 2.1. Suppression of the *ppi1* phenotype by *sp2* mutations. (A) Representative images of 28-day-old plants. (B) Representative images of 10-day-old plants showing phenotypic differences between the different *sp2* alleles. (C) Quantification of chlorophyll extracted with DMF from whole 10-day-old

seedlings. Error bars represent SEM ($n = 4$). The p values were derived from t -tests of the means of the indicated genotypes.

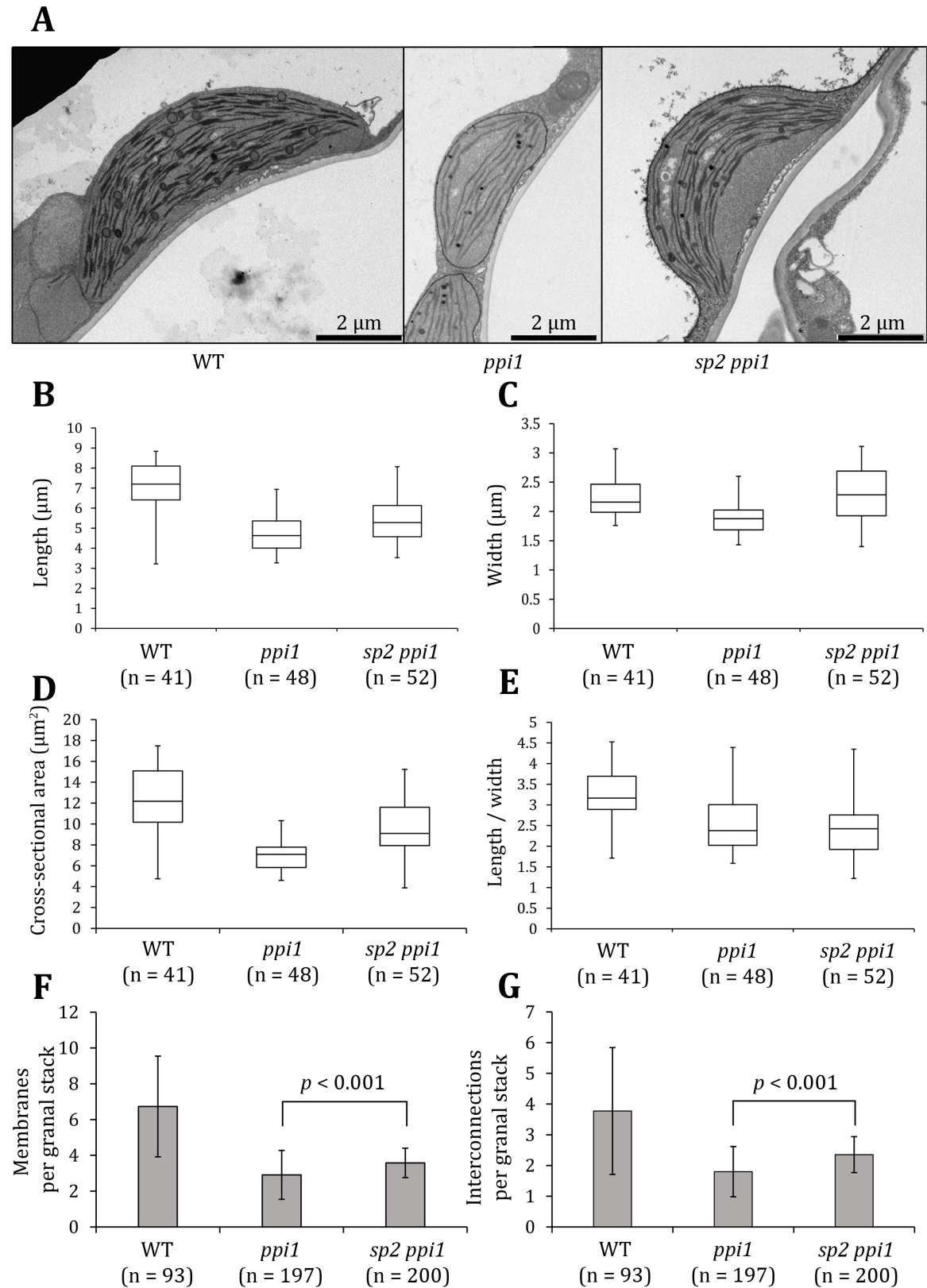


Figure 2.2. Chloroplast ultrastructure in *sp2 ppi1* mutants. (A) Representative transmission electron micrographs of chloroplasts in wild type, *ppi1* and *sp2 ppi1*. Black bars represent 2 μm. (B-E) Box and whisker plot comparisons of the length

(B) and width (C) of chloroplast cross-sections, chloroplast cross-section plan area (D), calculated as: $\pi \frac{(length \times width)}{2}$, and the chloroplast cross-section length/width ratio (E). Box and whisker plots represent median, median of upper quartile, median of lower quartile, and range of data. (F) Quantifications of thylakoid structure: the number of lamellae per granal stack, and the number of membrane interconnections per granal stack, was determined for each genotype. Error bars represent SD. The *p* values were derived from t-tests of the means of the indicated genotypes. Numbers of measurements (n) are shown in the figure. Measurements were taken across three individuals of each genotype.

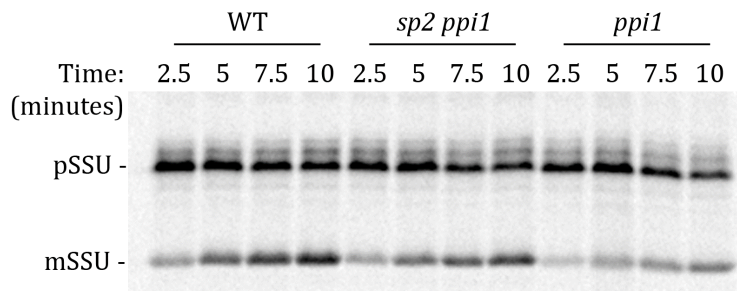
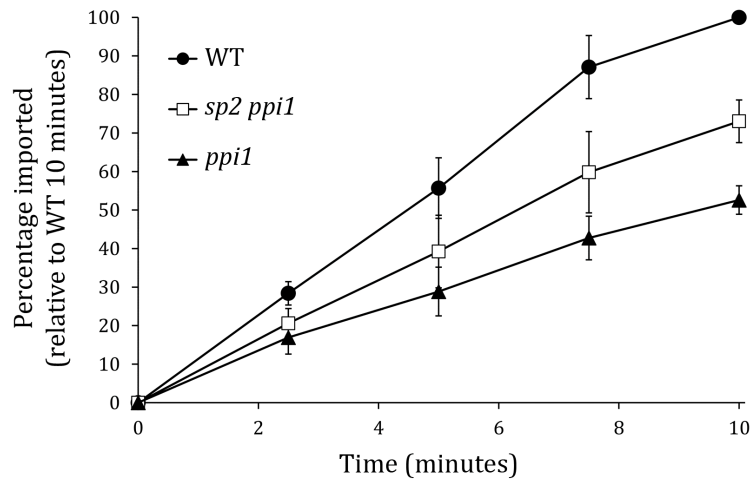
A**B**

Figure 2.3. Chloroplast protein import efficiency in *sp2 ppi1*. Radio-labelled pre-RUBISCO small subunit was incubated with chloroplasts isolated from wild-type, *ppi1* and *sp2 ppi1* plants over a 10-minute time-course to assay the efficiency of import and processing of preprotein substrates into mature chloroplast proteins. (A) A representative phosphorimage of samples taken throughout the import assay, each sample containing an equal volume of the import reaction, containing 10^7 chloroplasts. pSSU = pre-RUBISCO small subunit; mSSU = mature RUBISCO small subunit. (B) Quantification of the accumulation of mSSU in each genotype over time. Values shown are means derived quantification of phosphorimages of mSSU from three independent

experiments. All data were expressed relative to the value for wild type chloroplasts at 10 minutes. Error bars represent SEM ($n = 3$).

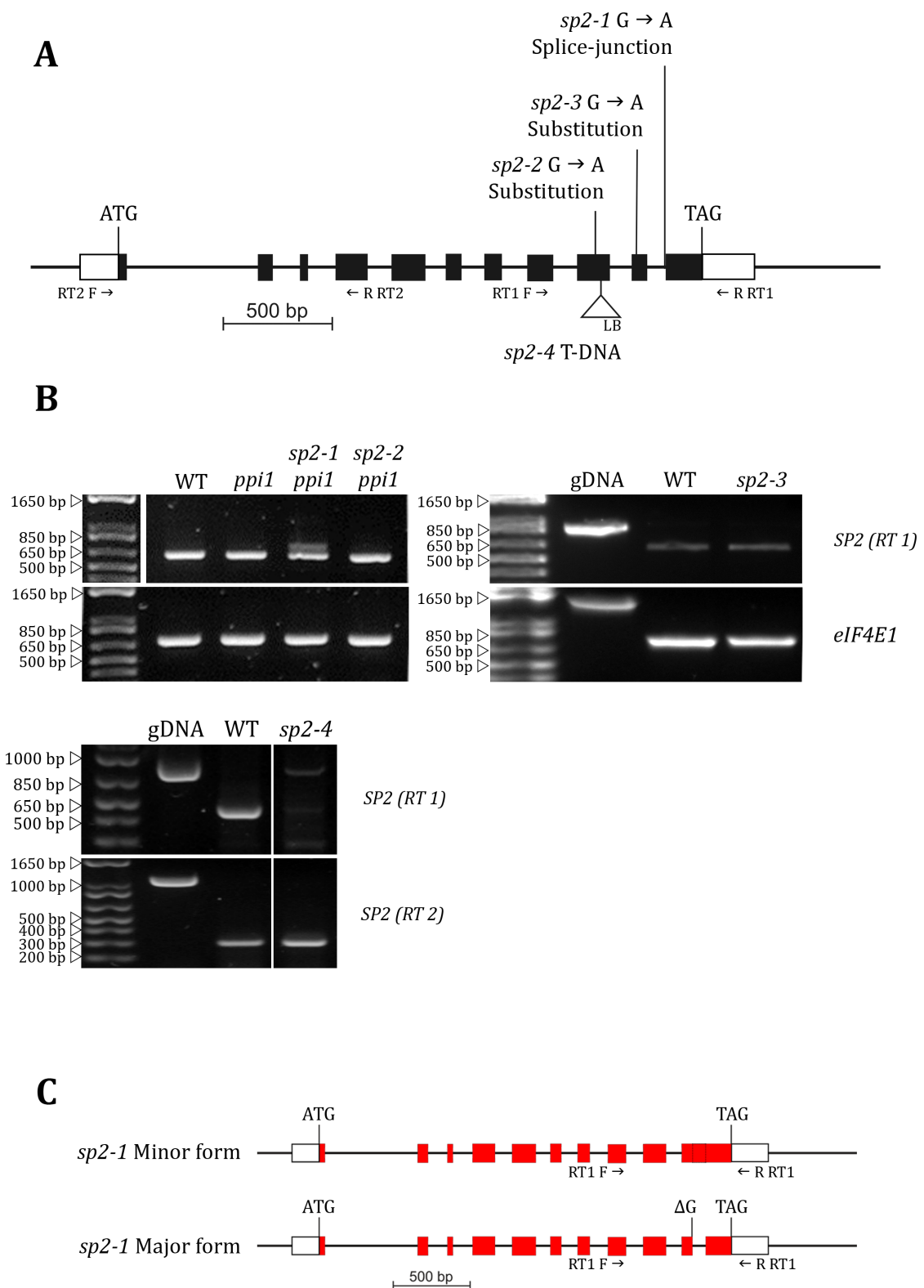


Figure 2.4. Analysis of *SP2* expression in the *sp2* mutants. (A) Schematic representation of the At3g44160 genomic locus annotated with the positions of the *sp2* mutations. The predicted nature of each mutation is shown, as are the

positions of the PCR primer binding sites (arrows). Exons = black boxes; introns = black lines; untranslated regions = white boxes; LB = left border sequence of the SALK T-DNA. (B) RT-PCR analysis of cDNA from the *sp2* mutants, using the primer pairs shown in A. gDNA = genomic DNA. (C) Gene-model representation of the two *sp2-1* transcripts, informed by data from the Sanger sequencing of corresponding RT-PCR products. Mature cDNA = red boxes; introns = black lines; untranslated regions = white boxes; RT 1 = *SP2* RT-PCR primer pair 1; RT 2 = *SP2* RT-PCR primer pair 2; Δ G = deletion of a single guanosine nucleotide.

<i>SP2:</i>	1	MGAQKSIHAGRAKIDVNVDFTHKLCTSLMFPAFRDTSSPLSLVIGSLCIKHPNLFGGSEK
<i>sp2-1 Δ:</i>	1	MGAQKSIHAGRAKIDVNVDFTHKLCTSLMFPAFRDTSSPLSLVIGSLCIKHPNLFGGSEK
<i>sp2-1 Ins:</i>	1	MGAQKSIHAGRAKIDVNVDFTHKLCTSLMFPAFRDTSSPLSLVIGSLCIKHPNLFGGSEK
<i>sp2-2:</i>	1	MGAQKSIHAGRAKIDVNVDFTHKLCTSLMFPAFRDTSSPLSLVIGSLCIKHPNLFGGSEK
<i>sp2-3:</i>	1	MGAQKSIHAGRAKIDVNVDFTHKLCTSLMFPAFRDTSSPLSLVIGSLCIKHPNLFGGSEK
<i>sp2-5:</i>	1	MGAQKSIHA--AKIDVNVDFTHKLCTSLMFPAFRDTSSPLSLVIGSLCIKHPNLFGGSEK
<i>SP2:</i>	61	LDVSWDKGLYDSNVLVAFRRRPRPEWRPQQCFFIQHSLSPEIGVHGTPVDNFSRSGSGGVN
<i>sp2-1 Δ:</i>	61	LDVSWDKGLYDSNVLVAFRRRPRPEWRPQQCFFIQHSLSPEIGVHGTPVDNFSRSGSGGVN
<i>sp2-1 Ins:</i>	61	LDVSWDKGLYDSNVLVAFRRRPRPEWRPQQCFFIQHSLSPEIGVHGTPVDNFSRSGSGGVN
<i>sp2-2:</i>	61	LDVSWDKGLYDSNVLVAFRRRPRPEWRPQQCFFIQHSLSPEIGVHGTPVDNFSRSGSGGVN
<i>sp2-3:</i>	61	LDVSWDKGLYDSNVLVAFRRRPRPEWRPQQCFFIQHSLSPEIGVHGTPVDNFSRSGSGGVN
<i>sp2-5:</i>	59	LDVSWDKGLYDSNVLVAFRRRPRPEWRPQQCFFIQHSLSPEIGVHGTPVDNFSRSGSGGVN
<i>SP2:</i>	121	LSKLALGLDLSEPASSKWSSTTSIKFEHVRPINDDGRAITRDLDGFPITCSGNTHDSMVV
<i>sp2-1 Δ:</i>	121	LSKLALGLDLSEPASSKWSSTTSIKFEHVRPINDDGRAITRDLDGFPITCSGNTHDSMVV
<i>sp2-1 Ins:</i>	121	LSKLALGLDLSEPASSKWSSTTSIKFEHVRPINDDGRAITRDLDGFPITCSGNTHDSMVV
<i>sp2-2:</i>	121	LSKLALGLDLSEPASSKWSSTTSIKFEHVRPINDDGRAITRDLDGFPITCSGNTHDSMVV
<i>sp2-3:</i>	121	LSKLALGLDLSEPASSKWSSTTSIKFEHVRPINDDGRAITRDLDGFPITCSGNTHDSMVV
<i>sp2-5:</i>	119	LSKLALGLDLSEPASSKWSSTTSIKFEHVRPINDDGRAITRDLDGFPITCSGNTHDSMVV
<i>SP2:</i>	181	LKQESRFAKATDQGLSHFSMQIEQGIPVVSQWLIENRKFVASKGVRFPAFLASLTGG
<i>sp2-1 Δ:</i>	181	LKQESRFAKATDQGLSHFSMQIEQGIPVVSQWLIENRKFVASKGVRFPAFLASLTGG
<i>sp2-1 Ins:</i>	181	LKQESRFAKATDQGLSHFSMQIEQGIPVVSQWLIENRKFVASKGVRFPAFLASLTGG
<i>sp2-2:</i>	181	LKQESRFAKATDQGLSHFSMQIEQGIPVVSQWLIENRKFVASKGVRFPAFLASLTGG
<i>sp2-3:</i>	181	LKQESRFAKATDQGLSHFSMQIEQGIPVVSQWLIENRKFVASKGVRFPAFLASLTGG
<i>sp2-5:</i>	179	LKQESRFAKATDQGLSHFSMQIEQGIPVVSQWLIENRKFVASKGVRFPAFLASLTGG
<i>SP2:</i>	241	SIVGDMAPYQAFaIGGLGsvRGYGEgAVGSGRSCLVAnTELALPLNKMTEGTIFLDcGTD
<i>sp2-1 Δ:</i>	241	SIVGDMAPYQAFaIGGLGsvRGYGEgAVGSGRSCLVAnTELALPLNKMTEGTIFLDcGTD
<i>sp2-1 Ins:</i>	241	SIVGDMAPYQAFaIGGLGsvRGYGEgAVGSGRSCLVAnTELALPLNKMTEGTIFLDcGTD
<i>sp2-2:</i>	241	SIVGDMAPYQAFaIGGLGsvRGYGEgAVGSGRSCLVAnTELALPLNKMTEGTIFLDcGTD
<i>sp2-3:</i>	241	SIVGDMAPYQAFaIGGLGsvRGYGEgAVGSGRSCLVAnTELALPLNKMTEGTIFLDcGTD
<i>sp2-5:</i>	239	SIVGDMAPYQAFaIGGLGsvRGYGEgAVGSGRSCLVAnTELALPLNKMTEGTIFLDcGTD
<i>SP2:</i>	301	LGSSRLVPGNPSMRQKPGFGYGFYGLRFKSPLGHLQVDYAINAFNQKTLyFGVTNLAS
<i>sp2-1 Δ:</i>	301	LGSSRLVPEtLQ-----
<i>sp2-1 Ins:</i>	301	LGSSRLVPGKFTLPSPNKNHKEVETQK-----
<i>sp2-2:</i>	301	LGSSRLVPGNPSMRQKPGFGYGFYGLRFKSPLGHLQVDYAINAFNQKTLyFGVTNLAS
<i>sp2-3:</i>	301	LGSSRLVPGNPSMRQKPGFGYGFYGLRFKSPLGHLQVDYAINAFNQKTLyFGVTNLAS
<i>sp2-5:</i>	299	LGSSRLVPGNPSMRQKPGFGYGFYGLRFKSPLGHLQVDYAINAFNQKTLyFGVTNLAS
<i>SP2:</i>	361	ST
<i>sp2-1 Δ:</i>	--	--
<i>sp2-1 Ins:</i>	--	--
<i>sp2-2:</i>	361	ST
<i>sp2-3:</i>	361	ST
<i>sp2-5:</i>	359	ST

Figure 2.5. Alignment of the predicted amino-acid sequences encoded by the *sp2* mutant alleles. Predicted protein sequences were informed by data from the Sanger sequencing of amplified *SP2* cDNA from each allele. Data referring to both transcripts from *sp2-1* are shown. *sp2-1 Δ* = *sp2-1* major transcript with nucleotide G deletion in the exon 10; *sp2-1 Ins* = *sp2-1* minor transcript, which retains part of intron 10.

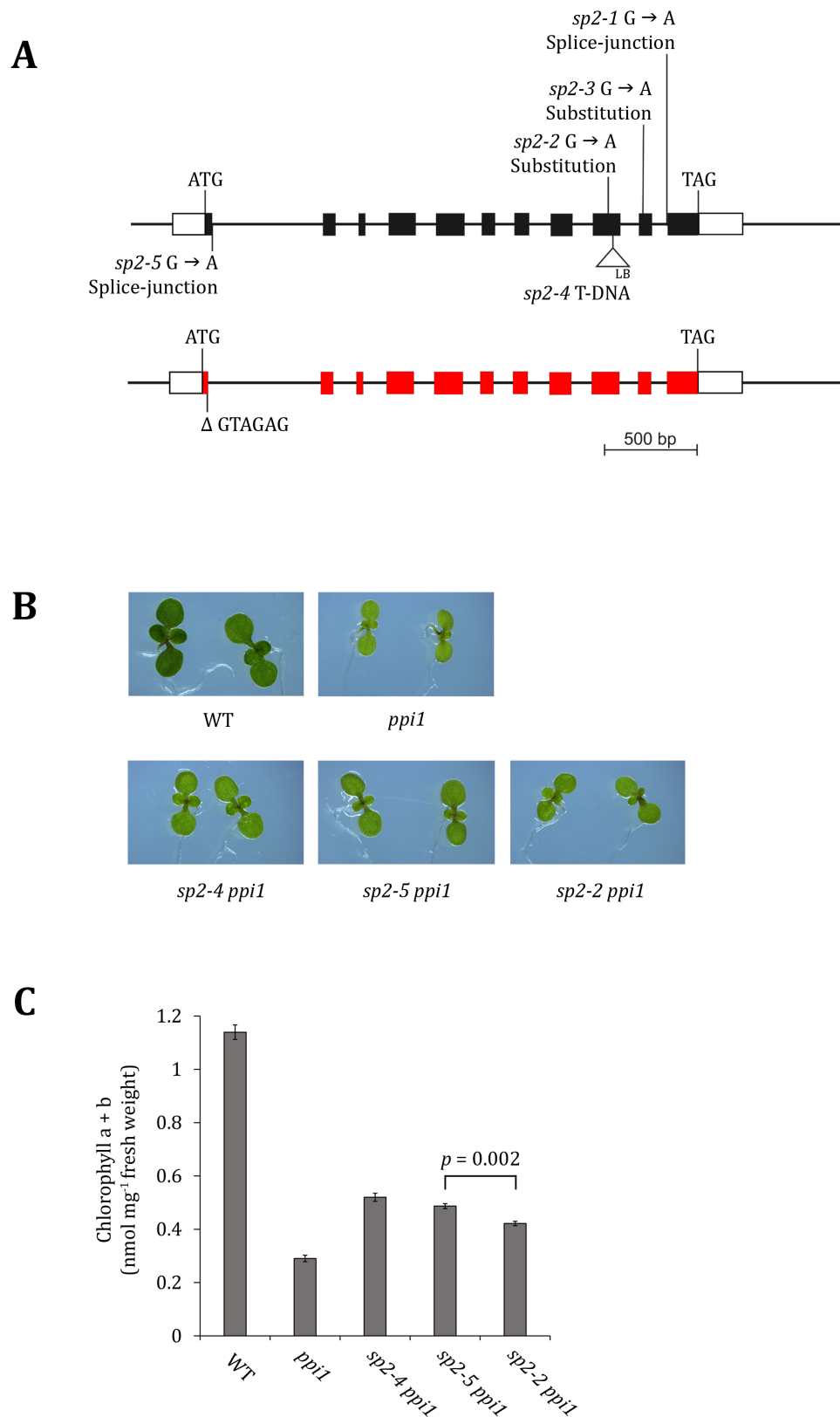


Figure 2.6. Analysis of the *sp2-5* mutant. (A) Schematic representation of the *SP2* genomic locus (top) and a gene-model representation of the *sp2-5* transcript (bottom), showing position of the *sp2-5* mutation relative to the original mutations, and its predicted consequences. Exons = black boxes; transcribed

regions = red boxes; introns = black lines; untranslated regions = white boxes; Δ = deletion; LB = left border sequence of the SALK T-DNA. (B) Representative images of 10-day-old seedlings. (C) Quantification of chlorophyll extracted with DMF from whole 10-day-old seedlings. The *p* value was derived from a t-test of the means of the indicated genotypes. Values shown are means; error bars indicate SEM (n = 4).

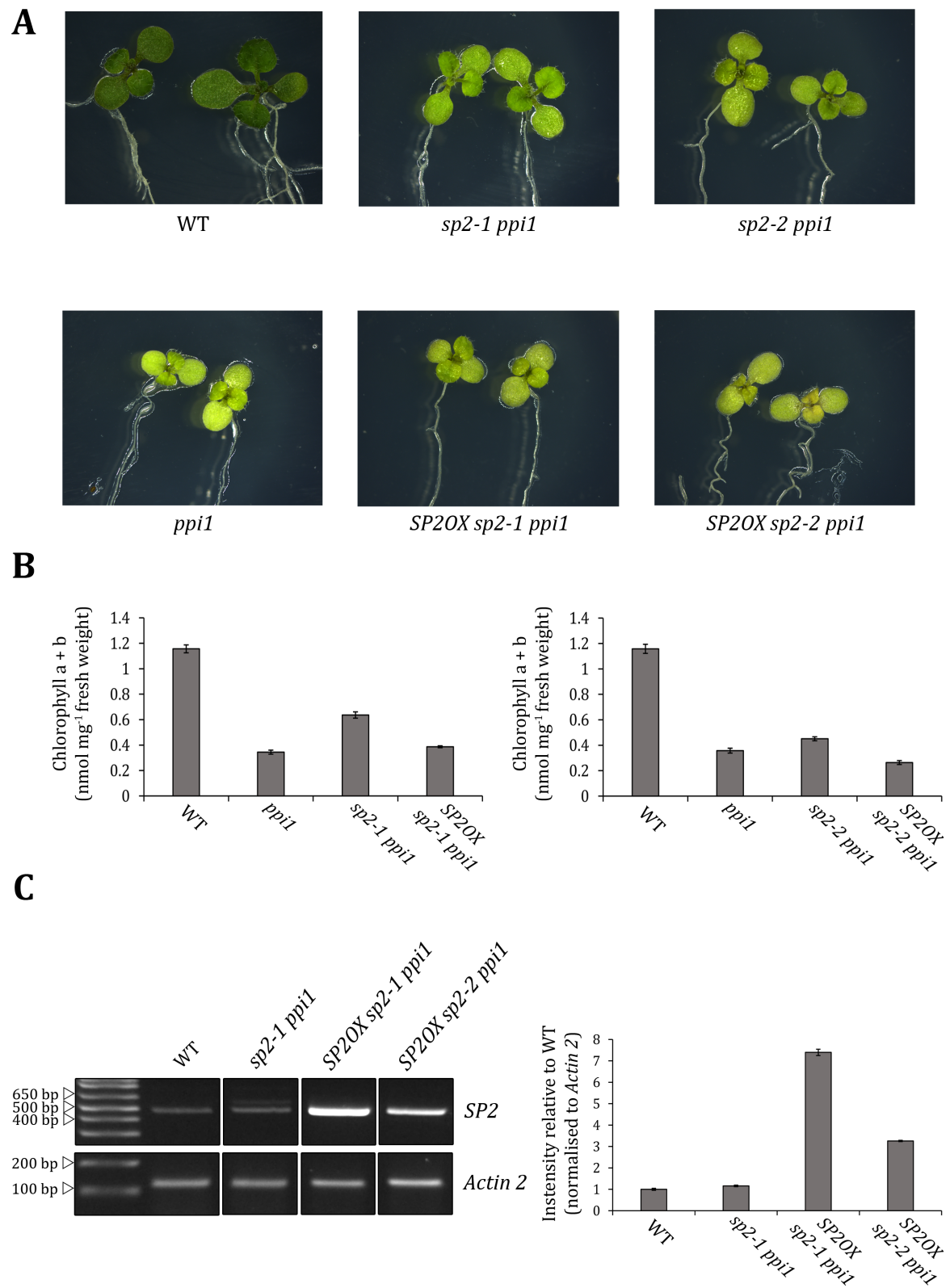


Figure 2.7. Complementation of *sp2 ppi1* mutants with the *SP2* CDS. (A) Representative images of 10-day-old seedlings. (B) Quantification of chlorophyll extracted with DMF from whole 10-day-old seedlings. The two graphs separately show the effects of complementation of *sp2-1* and *sp2-2* by *SP2OX*. (C) Semi-

quantitative RT-PCR analysis of *SP2* expression from *sp2 ppil* and *sp2 ppil SP2OX* plants, using *Actin 2* data as a normalisation control. Representative gel images are shown (left), alongside quantification derived from four biological replicates. Values shown in B and C are means; error bars indicate SEM (n = 4).

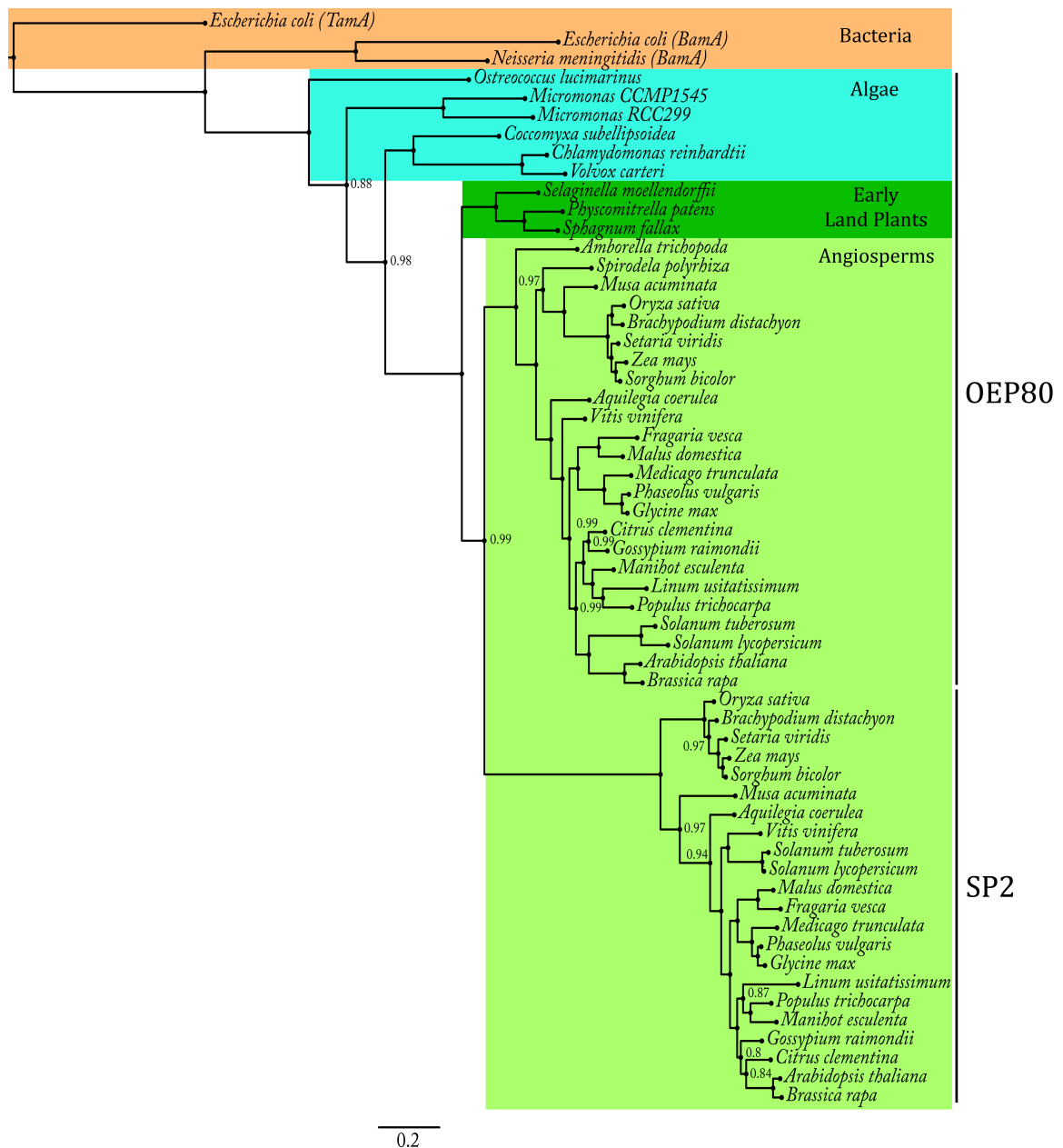


Figure 2.8. Bayesian inference phylogeny of SP2 and OEP80. Predicted protein sequences homologous to atSP2 and atOEP80 from a variety of plants and algae were retrieved from the Phytozome 12 database (Goodstein et al., 2012). Bacterial Omp85 sequences are included to act as an outgroup. Sequences were aligned with MAFFT (Kato and Standley, 2013), and the phylogeny was inferred using MrBayes 3.2 software (Ronquist et al., 2012). The tree represents a 50% consensus of 2000 trees generated from two runs, each using 8 Markov chain Monte-Carlo (MCMC) chains for 2 million generations. Only posterior probability values of < 1 are shown. Scale bar represents 0.2 changes per site.

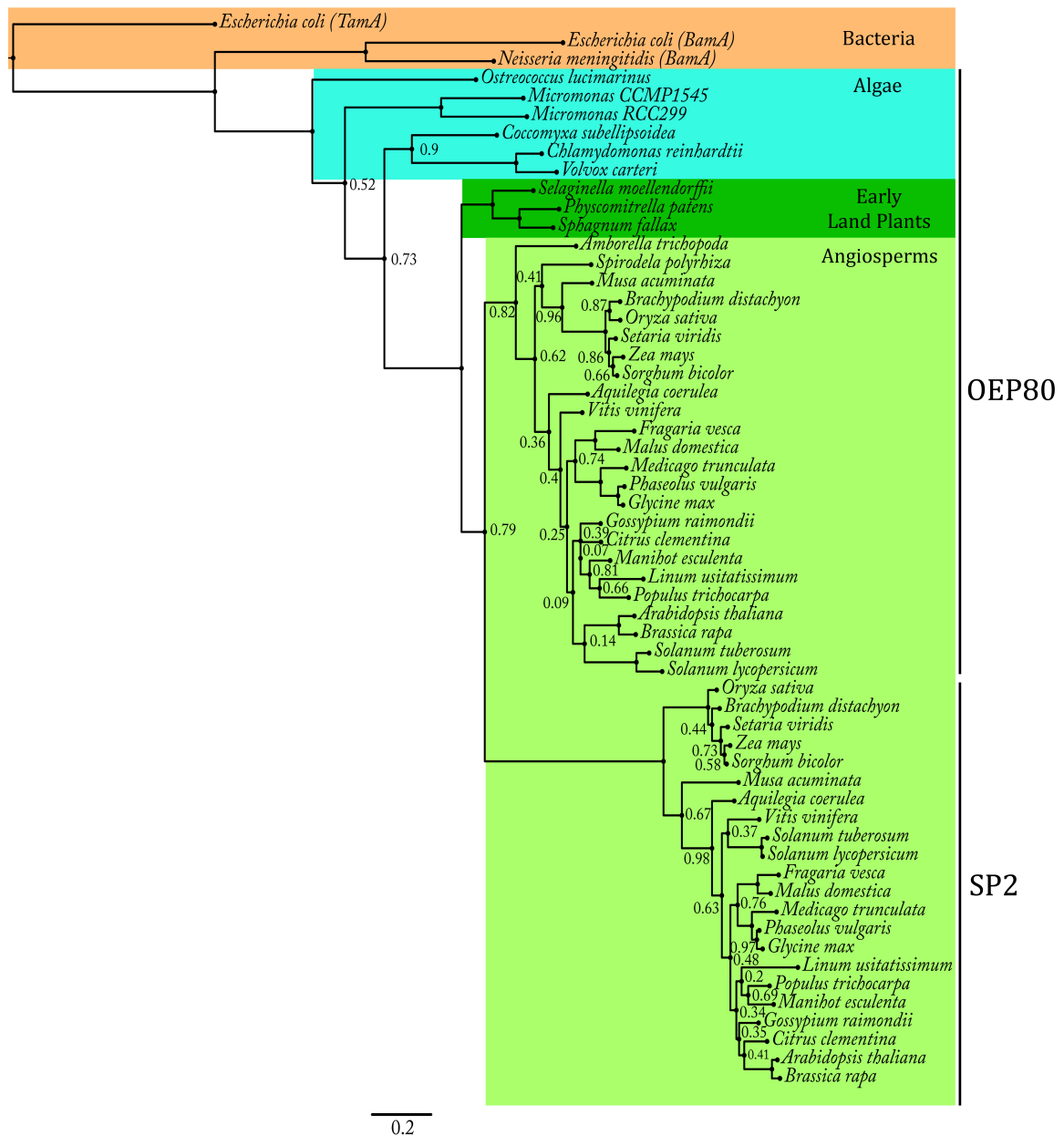


Figure 2.9. Boot-strapped maximum likelihood phylogeny of SP2 and OEP80. Predicted protein sequences homologous to atSP2 and atOEP80 from a variety of plants and algae were retrieved from the Phytozome 12 database (Goodstein et al., 2012). Bacterial Omp85 sequences were included to act as an outgroup. Sequences were aligned with MAFFT and the phylogeny was inferred using RAXML, and bootstrapped 1000 times (Stamatakis, 2014). Only boot-strap values of < 1 are shown. Scale bar represents 0.2 changes per site.

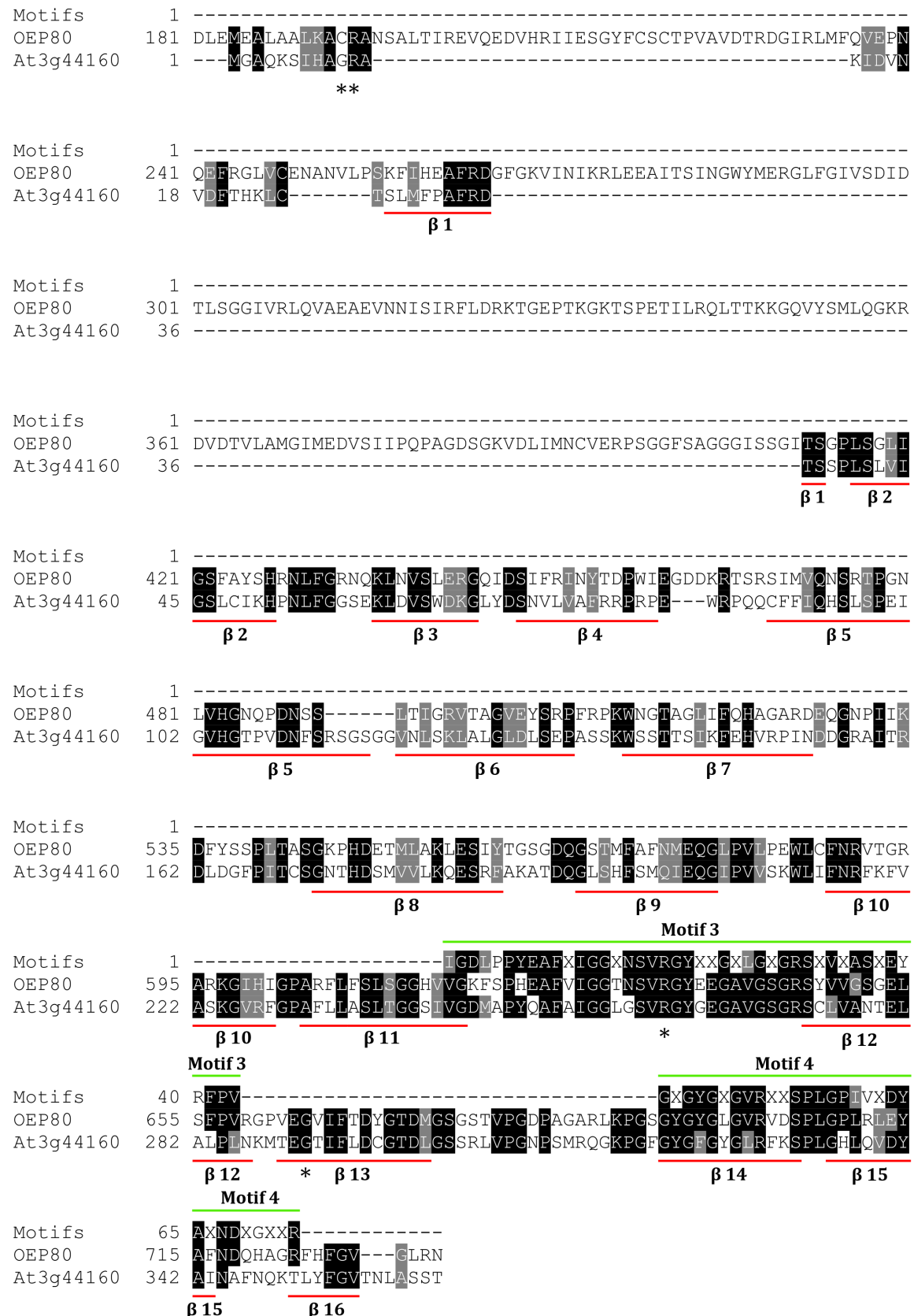


Figure 2.10. Alignment of the amino-acid sequences of SP2, OEP80, and conserved motifs 3 and 4 identified by Moslavac et al. (2005). The predicted positions of SP2 β -sheets are indicated by red lines. Positions of the conserved

motifs are indicated by green lines. Positions mutated in the *sp2* mutants are marked with asterisks. Amino-acid sequence identity between SP2 and OEP80 is 41%. Black = amino acids that are the same; grey = amino acids that share similar properties.

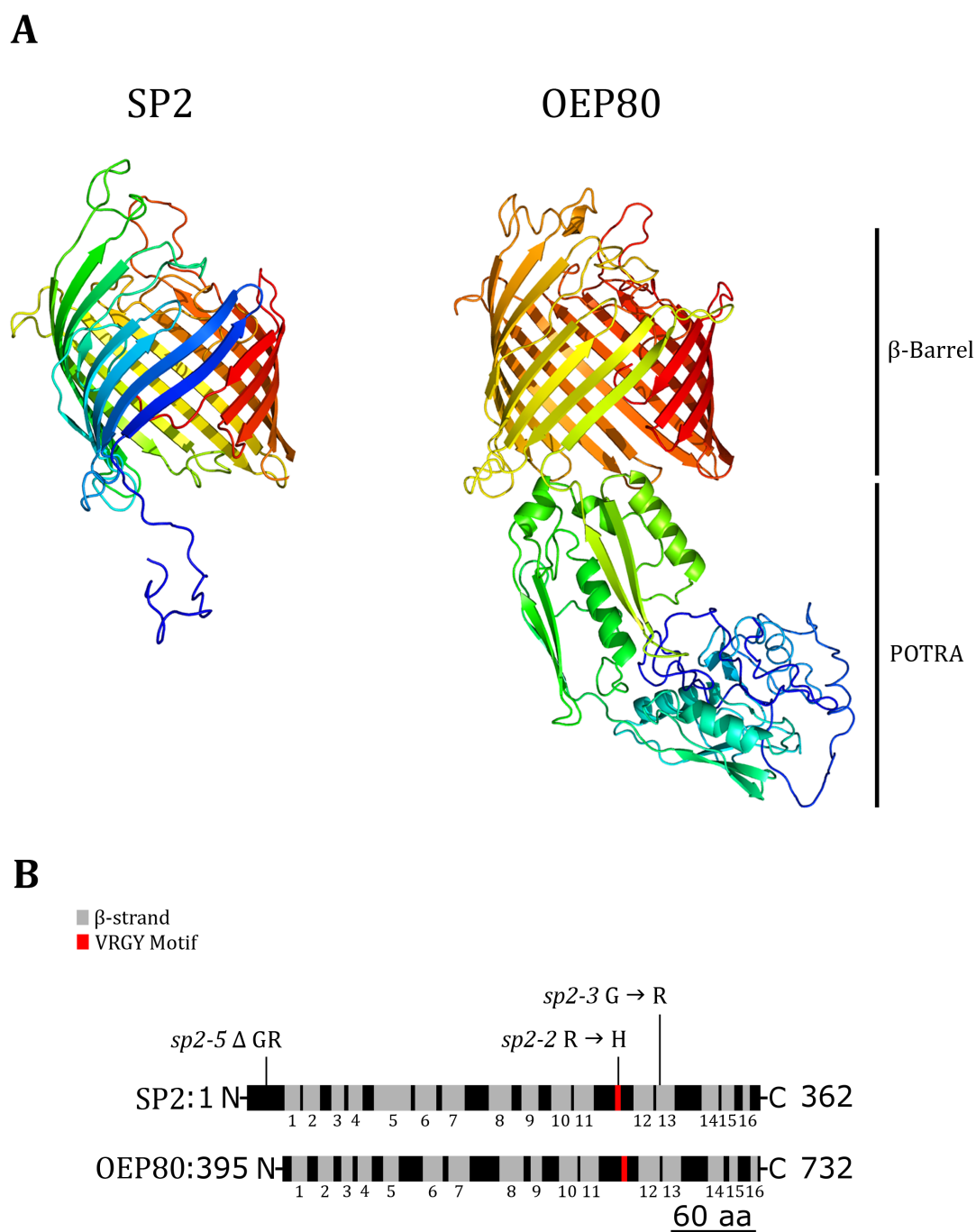


Figure 2.11. Structural models for the SP2 and OEP80 proteins. (A) Three-dimensional models for SP2 and OEP80 were derived by homology modelling using the crystal structure of *E. coli* BamA, using Swiss-model (Biasini et al., 2014). (B) Schematic diagrams showing secondary structure predictions for SP2 and the β-barrel portion of OEP80 (amino acids 395–732), annotated with the positions of the *sp2* mutations. Grey boxes = β-sheets; red box = VRGY motif. Scale bar represents 60 amino acids.

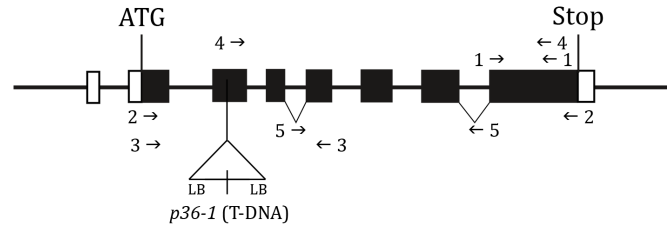
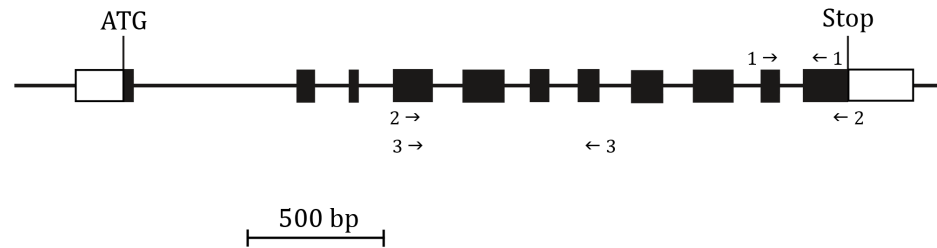
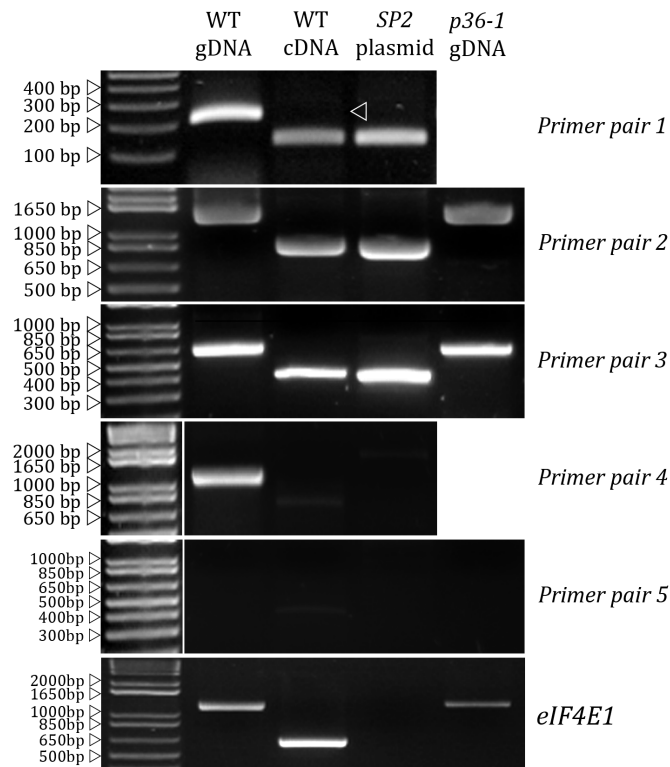
AAt3g48620 / *P36*At3g44160 / *SP2***B**

Figure 2.12. Analysis of the *P36* locus. (A) Schematic diagrams of the At3g48620/*P36* and At3g44160/*SP2* genomic loci. Annotations show the positions of the *P36-1* T-DNA insertion (Nicolaisen et al., 2015) and the RT-PCR primers used in B (arrows); note that each primer pair is numbered.

Exons = black boxes; introns = black lines; untranslated regions = white boxes; LB = left border sequence of the SALK T-DNA. (B) RT PCR analysis of the *P36* locus. Amplification with primer pair 1 (1). Amplification with primer pair 2 (2). Amplification with primer pair 3, which was the pair used for quantitative real-time PCR analysis of *P36* expression by Nicolaisen et al. (2015) (3). Amplification with primer pair 4 (4). Amplification with primer pair 5 (5). Amplification of the *eIF4E* control (6). gDNA = genomic DNA from wild type or the *P36-1* T-DNA insertion line; *SP2* plasmid = pDONR201 containing the *SP2* CDS sequence; bp = sizes of standards in nucleotide base-pairs.

Locus	Predicted to Encode	Predicted Localisation
Chromosome 1		
AT1G40104	Unknown protein	Mitochondria (TAIR), Cytosol (SUBA)
AT1G66730	DNA Ligase 6	Plastid (SUBA)
Chromosome 2		
AT2G01008	Unknown protein	Nucleus (SUBA)
AT2G07981	Unknown protein	Nucleus (SUBA)
AT2G08986	Unknown protein	Peroxisome (SUBA)
Chromosome 3		
AT3G43148	Unknown protein	Nucleus (SUBA)
AT3G43150	Unknown protein	Nucleus (SUBA)
AT3G44160	Omp85 family protein	Plastid (TAIR), Cytosol (SUBA)
AT3G44440	Unknown protein	Nucleus (SUBA)
Chromosome 4		
AT4G22200	Akt2, Potassium transporter 2/3	Cytosol (SUBA)
AT4G05612	Unknown protein	Cytosol (SUBA)
Chromosome 5		
AT5G12370	Sec10, Exocyst complex component	Plasma membrane (TAIR), Cytosol (SUBA)
AT5G52530	Dentin sialophosphoprotein-related	Nucleus (SUBA)

Table 1. Loci discovered to contain mutations in all three of the *sp2* alleles with whole genome sequences, compared to the *ppil-1* genetic background. These loci were identified by Dr Mats Töpel. TAIR = The Arabidopsis Information Resource (Berardini et al., 2015); SUBA = The SUBcellular localisation database for Arabidopsis proteins (Hooper et al., 2014).

Allele	Position Mutated in At3g44160	Nature of Mutation	No. of Reads with Mutation	Percentage of Reads with Mutation
<i>sp2-1</i>	2600	G → A	34	100
<i>sp2-2</i>	2246	G → A	26	100
<i>sp2-3</i>	2389	G → A	26	100

Table 2. Mutations discovered in the locus At3g44160 through the next-generation sequencing of three *sp2* alleles. Positions given are from the beginning of the At3g44160 annotation. No. of reads, indicates number of sequenced fragments spanning the locus of the mutation. Percentage of reads, indicates proportion of sequenced fragments spanning the locus of the mutation in which the mutation was observed. Arrow indicates base substitution.

Template	Primer Pair:				
	1	2	3	4	5
At3g48620 gDNA	1576	668	212	1088	-
SP2 gDNA	1579	674	210	-	-
At3g48620 cDNA	966	373	212	634	377
SP2 cDNA	905	390	125	-	-

Table 3. Predicted amplicon sizes in base-pairs, for five different primer pairs, designed to amplify the predicted At3g48620 cDNA. gDNA = genomic DNA.

Chapter 3

Results II

Functional analysis of SP2

Abstract

Understanding of the regulation of chloroplast biogenesis, and specifically the regulation of protein import into chloroplasts, is still underdeveloped. To address this deficiency, a mutagenesis screen was conducted for suppressors of chloroplast biogenesis mutant *ppi1*, a mutant of the Toc33 chloroplast protein import receptor. In Chapter 2, it was found that the identified *suppressor of ppi1 locus 2 (sp2)* is functionally related to chloroplast biogenesis. The SP2 protein is conserved in nearly all angiosperms, and is an Omp85 β -barrel protein closely related to OEP80. In this chapter, SP2 was further characterised. Genetic interaction experiments indicated that SP2 function has a specific relationship to protein import across the outer-envelope membrane of chloroplasts, and does not have a wider influence on chloroplast biogenesis. The effect of SP2 on the TOC proteins was studied at the molecular level by western blotting. This revealed that SP2 negatively influences the abundance of Toc33, Toc75-III and Toc159, and that these effects are delivered post-translationally. The mechanism of *ppi1* suppression by *sp2* was therefore determined to be related to the alleviation of negative regulation of these TOC proteins. Furthermore, co-immunoprecipitation assays demonstrated a specific interaction between SP2 and Toc33, Toc75-III, and Toc159. The effects of SP2 are similar to those of SP1, and the interactions between the two SP proteins will be studied in Chapter 4.

3. Functional analysis of SP2

3.1 Introduction

As a critical component of chloroplast biogenesis and a regulator of plastid function, the TOC complex has received a lot of attention. The core components were identified and defined long ago, but understanding of their regulation is lacking. To address this issue, a mutagenesis screen was performed to identify second-site suppressors of *ppi1*. The identification of the gene *SP2* as a result of this screen was described in Chapter 2. It was shown that SP2 is functionally linked to chloroplast biogenesis, and that it is related to the chloroplast outer-envelope component OEP80; thus, SP2 is β -barrel membrane protein, and a member of the Omp85 superfamily of proteins, which have functions in membrane protein biogenesis and protein transport across membranes. SP2 was shown to differ significantly from OEP80, and could represent a functionally distinct protein.

SP2 was revealed to lack POTRA domains, which are normally a critical feature of Omp85 proteins. These typically function in protein-protein interactions, as chaperones, and have displayed channel-gating activity. Such “POTRA-less” Omp85 proteins are rare, and have received limited characterisation. Similarly, OEP80 has not been much studied, and its proposed function in membrane protein biogenesis, though consistent with available evidence, remains speculative. Therefore, prior to this work there was little information regarding SP2 function, and so further characterisation was required to better understand SP2, its function, and its relationship to chloroplast biogenesis.

This chapter begins with genetic studies of *sp2* and *SP2* overexpressor (*SP2OX*) transgenics, and their genetic interactions with other mutants with defects related to chloroplast biogenesis. Subsequently, the molecular basis for phenotypic suppression of *ppi1* by *sp2* was investigated. Western blotting and qPCR were used to analyse the effects of *SP2* on specific chloroplast proteins. Finally, an epitope-tagged *SP2* construct was generated and used to assess *SP2* molecular localisation, topology and protein-protein interactions.

3.2 Results

3.2.1 SP2 specifically regulates the TOC components

To begin elucidating SP2 function, *sp2* alleles were crossed into the wild type background. As noted in Chapter 2, when linkage mapping of the *sp2* locus was attempted, the *sp2* mutants identified in the *suppressor of ppil* screen appeared to be genetically linked to the *ppil-1* T-DNA mutation, despite the two loci normally being located on different chromosomes. This was most likely due to a chromosome translocation in *ppil-1*, a common phenomenon in T-DNA mutagenized plants, with one study reporting a rate of 19 % (Clark and Krysan, 2010).

Linkage was also observed when attempting to cross the *sp2* alleles into the wild type background from *ppil-1*, as only wild type-like or *sp2 ppil*-like plants were found in the F₂ generation. The *sp2 ppil* plants were found to comprise approximately 1/4 of F₂ plants, the ratio expected for a single segregating locus. This observation was confirmed by genotyping, by which only wild type, *sp2 ppil*, or *sp2^{+/-} ppil^{+/-}* plants could be found.

In order to obtain the mutant *sp2* alleles isolated from *ppil-1* for use in genetic studies, the F₂ progeny of crosses between alleles of *sp2 ppil* and wild type were screened to find recombinant plants that were homozygous for the relevant *sp2* allele and hemizygous for the *ppil-1* T-DNA. Two hundred wild type-like F₂ plants were screened for each *sp2* allele, yielding between 1 and 3 individual *sp2 ppil^{+/-}* recombinant plants per *sp2* allele. This can be considered a recombination frequency of roughly 1% (1 centimorgan), which suggests that *sp2* and *ppil-1* were within approximately 200 – 250 kb of each other on

chromosome 1 or 3 (Lamesch et al., 2012). Subsequently, plants homozygous for each *sp2* allele and possessing two functional copies of the *PPII* (*TOC33*) gene were screened for, and identified, in the F3 generation.

3.2.1.1 Mutation or overexpression of *SP2* in the wild type genetic background produces no visible phenotype

In the wild type background (as opposed to *ppi1* mutant background in which they were first identified), the *sp2-1* to *-4* mutants all displayed no visible abnormal phenotypes as compared to the wild type (Fig. 3.1). Consistent with this observation, there was no significant difference in chlorophyll content between these four alleles and wild type (Fig. 3.1).

As no phenotype was observed when *sp2* was mutated, the effect of overexpression of *SP2* was tested, to see whether in this scenario there would be a phenotype which would be informative in determining *SP2* function. Wild type plants were transformed with the *SP2OX* construct, and lines overexpressing the *SP2* transcript were identified (Fig. 3.2). The two chosen lines, #12 and #15, were identified as segregating in a 3:1 ratio for phosphinothricin (PPT) resistance/sensitivity in the T2 generation, indicating that each possessed a single T-DNA insertion locus. In the T3 generation, these lines were identified to have among the highest levels of *SP2* expression by semi-quantitative RT-PCR, and so were chosen for further study.

Like the *sp2* mutants, these plants displayed no visible abnormal phenotypes, and no apparent differences in chlorophyll content relative to wild type (Fig. 3.2). Therefore, further experimentation was pursued to characterise the role of *SP2*.

3.2.1.2 The phenotype suppression effect conferred by *sp2* is specific to mutations of genes encoding TOC proteins

As mutations of *sp2* can suppress the phenotype of *ppi1*, a mutant of Toc33 which displays chloroplast biogenesis defects, it was important to know if *sp2* has a specific role in the regulation of plastid protein import or a more general effect on chloroplast biogenesis. To address this question, *sp2* mutants were crossed to other chloroplast protein import mutants.

Toc75-III is the main protein-conducting channel of the TOC complex. Complete knockout of Toc75-III is embryo lethal (Baldwin et al., 2005), but hypomorphic alleles of Toc75-III which exhibit chlorotic phenotypes have been identified, including *toc75-III-3* (Stanga et al., 2009; Huang et al., 2011). The *sp2* mutant, *sp2-4*, was able to suppress the chlorotic phenotype of *toc75-III-3* in the double homozygous state (Fig. 3.3A, C). These plants are both wild-type in appearance and contain chlorophyll concentrations comparable to those in wild type plants (Fig. 3.3C).

Toc159 is a major receptor of the TOC complex, and responsible for importing photosynthesis-related proteins (Bauer et al., 2000; Jarvis, 2008). Total knockouts of Toc159 are albino and lack development of the thylakoid membranes (Bauer et al., 2000). Therefore, we would not necessarily expect *sp2* to be able to recover chloroplast development in a total knockout of Toc159. However, a hypomorphic allele of Toc159 (*ferrochelatase two suppressor locus 1 [fts-1]*; here termed *ppi2-3*) was recently identified and found to be useful for this study (Woodson et al., 2015). The mutant *ppi2-3*, is caused by the mutation Q1272stop, leading to truncation of its final 12 amino acids, which are in the membrane anchored domain (M-domain) of Toc159. As with *ppi1* and *toc75-III-*

3, *sp2-4* can significantly suppress the *ppi2-3* phenotype (Fig. 3.3B, D). The double-mutant plants appeared both larger and greener (Fig. 3.3B), and displayed a significant increase in chlorophyll content (Fig. 3.3D).

As mutation of *SP2* causes the recovery of chloroplast biogenesis defects, it was hypothesized that overexpression of *SP2* would accentuate the effects of such defects. To test this, the *SP2OX #12* line was crossed to *ppi1-2* and *ppi2-3*. As expected, the phenotype of the *SP2OX ppi2-3* plants was more severe than that of either *ppi2-3* or *SP2OX #12* plants independently (Fig. 3.4A). This was also the case for *SP2OX ppi1* plants, both chlorophyll content, and rosette diameter, measured at its widest point were reduced with respect to *ppi1* plants (Fig. 3.4A, B). These phenotypic effects of *SP2OX* on *ppi1* plants are more subtle than those of *SP2OX* on *ppi2-3*, despite clear overexpression of *SP2* in *SP2OX ppi1* plants (Fig. 3.4C); the reasons for this are as yet unclear.

Next, to test if the *SP2*-mediated regulation of chloroplast biogenesis is specific to protein transport across the OEM, *sp2* mutants were crossed to mutants of other plastid protein import components which are not localised to the OEM.

Tic22 is a chloroplast protein which is present in organisms across the Viridiplantae, and has homologues in extant cyanobacteria (Kasmati et al., 2013). Mutants of Tic22 have a chloroplast biogenesis defect, which is specifically related to protein import (Kasmati et al., 2013; Rudolf et al., 2013). Importantly, it has been speculated that the relationship of Tic22 to plastid protein import could be via a role in the assembly of the Toc75 and OEP80 proteins. This was proposed due to an apparent similarity in its tertiary structure to *E. coli* BamB, which is involved in the biogenesis of bacterial β -barrel outer-membrane proteins

(OMPs; Brouwer et al., 2012). Because SP2 is a close relative of OEP80, and because OEP80 is hypothesised to have a role in assembling β -barrel OMPs like Toc75 (Huang et al., 2011) like its bacterial relative *E. coli* BamA (Misra, 2012; Noinaj et al., 2015), it could be that the function of SP2 is to negatively regulate the function of OEP80. Thus, it was important to test the effects of *sp2* on mutants of *tic22*, a hypothesised component of β -barrel OMP assembly in chloroplasts.

Arabidopsis has two isoforms of Tic22, which are both localised to the intermembrane space and have been shown to interact with the POTRA domains of Toc75-III (Kasmati et al., 2013; Paila et al., 2016). Individually, the mutants affecting these isoforms have only subtle phenotypes, but a defect is clearly seen in the double mutant, *tic22-III tic22-IV*, but even this is only apparent early in development (Kasmati et al., 2013; Rudolf et al., 2013). Thus, the double mutant was chosen for analysis in this study.

As well as an overall paleness due to a general lack of chlorophyll, *tic22-III tic22-IV* plants exhibit a distinct pale border at the basal margins of the first true leaves (Fig. 3.5A). The *sp2-4 tic22-III tic22-IV* triple mutant still exhibited the chlorotic phenotype and pale border at the edges of the first true leaves (Fig. 3.5A). While there was possibly a slight increase in chlorophyll content in *sp2-4 tic22-III tic22-IV* as compared with *tic22-III tic22-IV*, this difference was not statistically significant (Fig. 3.5D). Taken together, these results suggest that the *tic22-III tic22-IV* double mutant is not suppressed by *sp2*.

The *sp2* mutants were also crossed to a mutant of an inner-envelope protein import component, Tic40, and to a mutant of a stromal chaperone, Hsp93-V, and corresponding double mutants were identified. Mutants of Tic40

exhibit a chlorotic phenotype comparable to *ppi1*, which is linked to a defect in plastid protein import (Kovacheva et al., 2005). Mutants of Hsp93-V have a mild chlorotic phenotype which also exhibits a plastid protein import defect (Kovacheva et al., 2005; Kovacheva et al., 2007).

The *sp2-4* mutation did not suppress the visible *tic40* phenotype, nor did it lead to increased chlorophyll content in the corresponding double mutant (Fig. 3.5B, E). Similarly, the *sp2-1* mutation did not suppress the visible *hsp93-V* phenotype, or lead to an increase in chlorophyll content in the double mutant (Fig. 3.5C, F).

3.2.1.3 SP2 affects the abundance of TOC proteins

It was important to discover the underlying mechanism for the suppression effect conferred by the *sp2* mutations on the chlorotic mutants of TOC proteins. As the recovery of the *ppi1* phenotype had been related to chloroplast protein import, assessing the abundance of the TOC proteins themselves was an obvious starting point.

First, the three alleles *sp2-2*, *sp2-3*, and *sp2-4* were analysed in the *ppi1* background. Western blots revealed a strong and reproducible increase in Toc75 abundance in the double mutants as compared to *ppi1* (Fig. 3.6A, B). In contrast, there were no clearly observable changes in the levels of Toc34, Toc159, or the nuclear protein H3, as compared to *ppi1* (Fig. 3.6A, B). Analysis by quantitative real-time PCR (qPCR) revealed no apparent changes in *TOC75* or *TOC159* mRNA expression, indicating that the observed effects on Toc75 protein abundance were not due to a transcriptional effect (Fig. 3.6C). The abundance of Toc75 in the double mutants was proportional to the observed suppression, as the increase in

Toc75 abundance was much less in *sp2-2 ppil*, which is hypomorphic and the weakest of the *sp2* alleles.

Similar results were obtained in the *toc75-III-3* mutant background. Toc75 abundance is strongly reduced in the *toc75-III-3* mutant allele, which may be due to the single amino-acid substitution in this allele resulting in a less stable protein. In *sp2-4 toc75-III-3* double mutants, however, there was a clear increase in the abundance of Toc75, along with that of Toc159, returning both to levels comparable with those in the wild type (Fig. 3.7). These data were also reflective of the visible phenotype of the relevant double mutant, which appears very similar to wild type (Fig. 3.3A).

Next, the single mutants, *sp2-2*, *sp2-3*, and *sp2-4*, were analysed, and similar effects were observed. Though these mutants display no abnormal visible phenotypes (Fig. 3.1), the Toc75 and Toc33 proteins are nonetheless more abundant than in the wild type (Fig. 3.8A, B). However, as in the *ppil* background, no convincing effect of the *sp2* mutations on Toc159 abundance was detected (Fig. 3.8A, B). The effects on protein abundance were apparently specific to the TOC proteins linked to the import of photosynthesis-related preproteins. The levels of Toc34, OEP80, the IEM proteins Tic110 and Tic40, and the mitochondrial inner-membrane protein, Slp1, did not appear to be affected (Fig. 3.8A, B).

As the mutation of *SP2* causes an increase in the abundance of certain TOC proteins, one would expect overexpression of *SP2* to cause a decrease in their abundance. This possibility was investigated, and it was clearly observed to be the case for all of the TOC proteins involved in the import of photosynthesis-related preproteins: Toc75, Toc33 and Toc159 (Fig. 3.8A, B). This effect was specific

because, as in the case of the *sp2* mutants, Toc34, OEP80, the IEM proteins Tic110 and Tic40, and the mitochondrial inner-membrane protein, Slp1, did not appear to be affected (Fig. 3.8A, B). Once again, qPCR analysis did not reveal any significant changes in *TOC75*, *TOC159* or *TOC33* mRNA expression in the *SP2* overexpressor lines, indicating the effects on the Toc75, Toc159 and Toc33 proteins are not transcriptional (Fig. 3.8C).

3.2.2 Localisation and interactions of SP2

Given the observed effects of *SP2* on the abundance of TOC proteins, the nature of this interaction was studied further. As effects of *SP2* on the mRNA expression of TOC genes was ruled out, a post-translational interaction seemed likely. Therefore, co-immunoprecipitation studies, to assess for possible physical interactions between SP2 and TOC proteins, were pursued. Due to the lack of an SP2 antibody, an epitope tagged version of SP2 was used.

3.2.2.1 *SP2-Myc* can complement mutants of *SP2*

SP2 was translationally fused at its C-terminus with six repeats of the c-Myc epitope, and the resulting gene chimera was placed under the control of the cauliflower mosaic virus 35S promoter, a construct henceforth called *SP2-Myc*. An epitope tag can have a negative impact on a protein's targeting and function. Thus, the function of the fusion protein was tested by generating plants overexpressing *SP2-Myc* in the *sp2 ppi1* background, to assess for phenotypic complementation. If *SP2-Myc* retains its normal function, the phenotype of *SP2-Myc sp2 ppi1* plants was predicted become similar to that of *ppi1*. The results obtained were as expected, in the case of both *SP2-Myc sp2-1 ppi1* and *SP2-Myc sp2-2 ppi1* (Fig. 3.9). These plants appeared smaller and had a lower chlorophyll

content than *sp2-1 ppil* and *sp2-2 ppil*, and were similar to *ppil* (Fig. 3.9). These results therefore indicated that SP2-Myc retains SP2's normal function, and so the fusion protein is suitable for studies on SP2 function.

3.2.2.2 SP2-Myc is detectable by western blotting using c-Myc antibody

The rationale for adding an epitope tag is to make the protein of interest reactive to an antibody, which is the fundamental basis for analytical techniques like western blotting and immunoprecipitation. Therefore, a total-protein extract of 10-day-old plants expressing *SP2-Myc* was analysed by anti-c-Myc western blotting to test the system's ability to detect SP2-Myc, and to assess whether the protein is of the predicted molecular weight. For this test, and for subsequent experiments, *SP2-Myc* was also introduced into *sp2-4*, i.e., a plant which does not carry the *ppil* mutation. This is important for downstream experiments, so that SP2-Myc can be studied in wild type-like plants, and so that there is no risk of interference from native SP2, and so that interactions with the Toc33 protein may be assessed.

As SP2 is 39.1 kDa, and the tag comprising six repeats of the c-Myc epitope used is 10.5 kDa (Dubin et al., 2008), a product of 49.6 kDa was expected. When blotting for c-Myc, a double band was observed – the main band just above the 50 kDa marker, and a smaller one just below it – and these bands did not appear in protein samples from wild type plants (Fig. 3.10). These locations are approximately correct for a protein of 49.6 kDa. The reason for the appearance of two bands is probably because the band was effectively split into two by the highly-abundant RUBISCO large subunit, which is 53 kDa in Arabidopsis (Lamesch et al., 2012; Bracher et al., 2017). Alternatively, the cause could be

some truncation of the protein, translation initiation at a downstream ATG, or perhaps some form of post-translational modification or processing. Regardless, considering the results from the complementation experiment, it can be concluded that SP2-Myc is a functional protein.

3.2.2.3 SP2-Myc is imported and assembled into the OEM

Two previous studies have shown SP2 localisation to the chloroplast. The first showed that radiolabelled, in vitro-translated SP2 could be imported into isolated pea chloroplasts (Day et al., 2014). After import and washing with Na_2CO_3 , SP2 was found predominantly in the pellet fraction, as opposed to the soluble fraction, indicating that most of the imported protein was integrated into chloroplast membranes (Day et al., 2014). However, treatment of the chloroplasts post-import with proteases revealed a sensitivity of SP2 to both thermolysin, which cannot penetrate the outer-envelope membrane, and trypsin, which can penetrate the outer membrane to degrade proteins in the intermembrane space and inner-membrane (Day et al., 2014). This indicated that the imported SP2 was exposed at the OEM, where it would be accessible to both proteases. This is consistent with an OEM topology, but is in contrast with results obtained for Toc75-III, which exhibited only partial sensitivity to thermolysin, and for the POTRA-less Toc75-IV, which exhibited no sensitivity to thermolysin (Baldwin et al., 2005; Day et al., 2014); thus, one may call into question the correct import and assembly of SP2 in this experiment. The second study detected native SP2 using an antibody raised against SP2, and it was shown that SP2 abundance correlated with that of Toc75 in samples taken from subcellular fractionation of *Arabidopsis* tissues, indicating that SP2 is localised in chloroplasts as that is known to be the case for Toc75 (Nicolaisen et al., 2015).

Given the presence and size of the Myc epitope tag, it is important to assess its effects on SP2 protein localisation and topology, and whether these are consistent with previous reports. Moreover, use of the Myc tagged SP2 would enable a more direct assessment of SP2 localisation and topology, given that Day et al. (2014) determined localisation from in vitro import, and Nicolaisen et al. (2015) assessed SP2 localisation indirectly. Additionally, the location of the Myc tag can be inferred by protease treatment and so the orientation of the SP2-Myc protein could be inferred. This is possible because, as the SP2 protein contains an even number of β -strands, both termini of the protein will face either the cytosol or the intermembrane space. The topologies of related proteins, OEP80 and Toc75-III, would indicate that the termini of the SP2 protein are likely to face the intermembrane space, and so the Myc-tag should be insensitive to thermolysin (Dastvan et al., 2016; Paila et al., 2016).

To test whether SP2-Myc is integrated into the chloroplast membranes, chloroplasts isolated from *SP2-Myc sp2-4* plants were treated with Na_2CO_3 . Large amounts of SP2-Myc were found in both the soluble and pellet fractions (Fig. 3.11A, B). This was the case also for soluble proteins Hsp93 and 1-deoxy-D-xylulose reductoisomerase (DXR), but not for Toc75-III or any other integral membrane proteins, with the exception of Toc33 (Fig. 3.11A, B). This indicates that SP2-Myc is not fully integrated into the membrane. This result is similar to that Day et al. (2014), where some SP2 is seen in the soluble fraction. It could be the case that, when overexpressed, SP2-Myc is imported into the chloroplast but not fully assembled. Another possibility is that unique properties of SP2 in the membrane lend itself to being more easily extracted than other proteins. In support of this idea, the cellular protein Toc33 is also partially extracted, in this

experiment (Fig. 3.11A, B); this result was reproducible and was previously reported (May and Soll, 1998).

To further elucidate the topology of SP2-Myc, isolated chloroplasts from *SP2-Myc sp2-4* plants were subjected to protease treatment. SP2-Myc was partially sensitive to thermolysin, similarly to Toc75-III, whereas the receptor proteins Toc33 and Toc159 showed a much higher sensitivity to thermolysin, indicative of their prominent cytosolic domains (Fig. 3.12). Note that Toc159 was not completely sensitive to thermolysin, as previously observed (Hiltbrunner et al., 2001). The IEM proteins Tic110 and Tic40 and the stromal chaperone Hsp93 did not show sensitivity to thermolysin, as expected (Fig. 3.12). SP2-Myc was sensitive to trypsin treatment, similarly to Toc159, though it was clearly not as sensitive as the other OEM proteins, Toc33 and Toc75-III (Fig. 3.12). The Tic110 protein (which may have large IMS domains; Tsai et al., 2013) had a similar sensitivity to trypsin as SP2-Myc, whereas Tic40 (which is a single-span protein of the IEM that resides largely in the stroma; Chou et al., 2003; Bédard et al., 2007) and stromal Hsp93 exhibited little sensitivity to trypsin (Fig. 3.12).

The partial sensitivity of SP2-Myc to trypsin and thermolysin was not exactly the same as either Toc75-III or Tic110, but SP2-Myc did display similar sensitivity to the OEM protein Toc159 (Fig. 3.12). As SP2-Myc was partially sensitive to thermolysin, this may indicate that SP2-Myc is inserted into the OEM. However, because SP2-Myc was not as sensitive to trypsin as Toc75-III, and was more comparable to Tic110 in this regard (which may have prominent IMS domains; Tsai et al., 2013), and because it was considerably more trypsin-sensitive than Tic40 and Hsp93, the protein could possibly be considered to exist (partly) in the intermembrane space. Thus, a possible explanation is that SP2-

Myc may only be partially assembled, which would be consistent with the results of the alkaline extraction experiment, where SP2-Myc was present in the soluble and pellet fractions (Fig. 3.11A, B).

With regard to the membrane orientation of SP2, as SP2-Myc exhibits only a partial sensitivity to thermolysin it is unlikely that it is assuming a topology with the termini facing the cytosol, as in this case the c-Myc epitopes would be prone to cleavage (Fig. 3.12). Therefore, it is likely that the SP2 termini face the IMS, as would be predicted based on the topology of the homologous protein Toc75 (Chen et al., 2016; Paila et al., 2016).

On the basis of the information available, two interpretations are possible. SP2 resides in the OEM and can be imported even when overexpressed, but overexpression causes accumulation of SP2 transport intermediates in the IMS owing to the saturation of the assembly machinery. Alternatively, SP2 resides in the OEM and can be imported even when overexpressed, but the presence of a large (10.5 kDa) epitope tag causes accumulation of SP2 in the IMS due to interference with the SP2 assembly machinery. Nevertheless, the evidence suggests that SP2-Myc is at least partially assembled and functional.

3.2.2.4 SP2-Myc physically interacts with TOC components

As *SP2* interacts genetically with mutations affecting the TOC proteins, and influences the accumulation of TOC proteins, it was important to elucidate the nature of the SP2-TOC interaction at the protein level. Given that *SP2-Myc* could complement *sp2 ppil* mutants, and localisation studies indicated a chloroplast localisation for the SP2-Myc protein, co-immunoprecipitation (co-IP) studies using these plants were pursued. To avoid problems with potential

unassembled SP2-Myc, the isolated chloroplasts were lysed with hypotonic buffer, and protein for co-IP was extracted from the recovered membranes.

Using beads conjugated with the c-Myc antibody, protein extracted from isolated chloroplast membranes of *SP2-Myc sp2-4* plants was used for co-immunoprecipitation analysis, employing SP2-Myc as bait and the native TOC proteins as prey. In this experiment, SP2-Myc specifically pulled down the TOC proteins Toc33, Toc75-III and Toc159, none of which were similarly pulled down when protein from the wild type chloroplasts was applied to the beads (Fig. 3.13). However, if SP2 is not fully imported, this interaction could be criticised as non-specific or the consequence of partial import of SP2-Myc. Contrary to this criticism these interactions were found to be specific, as SP2-Myc did not pull down the chloroplast IEM proteins Tic110 or Tic40 (Fig. 3.13). Together, these data indicate a specific physical interaction between SP2 and the TOC proteins analysed, consistent with the genetic interaction data which indicated that SP2 exerts negative regulation over these TOC proteins.

3.3 Discussion

As elaborated in Chapter 2, SP2 has a clear functional relationship to chloroplast biogenesis. In this chapter, it has been discovered that this relationship is specific to protein import across the chloroplast OEM via the TOC complex. Evidence for this link emerged from genetic interaction studies between *sp2* and various plastid protein import mutants (Fig. 3.3). Mutants of *sp2* could suppress the TOC mutants *ppi1*, *ppi2-3* and *toc75-III-3*, but not the intermembrane space component mutant *tic22-III tic22-IV*, the TIC component mutant *tic40*, or the stromal protein import chaperone mutant *hsp93-V* (Figs. 3.3, 3.5).

The mechanism of suppression of *ppi1* by *sp2* is linked to a Toc75-III protein abundance increase (Fig. 3.6), permitting enhanced import of preproteins into the plastid, which in turn enables further development of the chloroplast compared to *ppi1*. However, lack of functional Toc33 cannot be completely compensated for by the accumulation of Toc75-III, owing to the requirement for Toc33 for the import of certain photosynthesis-related preproteins (Kubis et al., 2003; Constan et al., 2004), and as a result the *sp2 ppi1* plants remain underdeveloped relative to the wild type. The suppression mechanism appears to be the same in the case of *sp2 toc75-III-3*, although in this case the plants do appear very similar to the wild type (Fig. 3.3). This is possible because the *toc75-III-3* mutation is hypomorphic and does not completely knockout Toc75-III, but instead reduces its abundance; thus, when the *sp2* mutation is introduced, the Toc75-III protein can accumulate resulting in complete suppression (Figs. 3.3, 3.7). This also suggests that the amino-acid substitution in the *toc75-*

III-3 mutant affects stability or assembly of the protein, as opposed to its protein transport function.

Characterisation of the *sp2* mutants and the *SP2* overexpressors revealed opposing effects on the TOC components Toc33, Toc159 and Toc75-III, at the protein level, albeit in the absence of visible phenotypic consequences (Figs. 3.1, 3.2, 3.8). Specifically, mutation of *SP2* causes an increase in the abundance of Toc33 and Toc75-III in the wild type background, of Toc75-III in the *ppi1* background, and of Toc159 and Toc75-III in the *toc75-III-3* background (Figs. 3.6, 3.7, 3.8). Conversely, overexpression of *SP2* causes the depletion of Toc33, Toc159 and Toc75-III in the wild type background. In both cases, these effects were specific and not accompanied by impacts on other plastid proteins or those in mitochondria or the nucleus (Fig. 3.8). Interestingly, no obvious impact on Toc34 abundance was observed in either mutants or overexpressors of *SP2*, suggesting that SP2 activity is specific to Toc75 and the TOC proteins linked to the import of photosynthesis-related preproteins (Figs. 3.6, 3.8).

Consistent with these observations, overexpression of *SP2* in the *sp2 ppi1* mutants caused reversion of the phenotype to that of *ppi1*, while in the *ppi2-3* background it caused accentuation of the mutant's chlorotic phenotype (Fig. 3.4). Curiously, there was only a mild phenotypic effect when *SP2OX* was introduced into the *ppi1* single-mutant background (Fig. 3.4). This may have to do with the nature of the *TOC159* hypomorphic mutation in *ppi2-3*, and the topology of the TOC complex. This issue will be revisited in Chapter 4.

These results are indicative of a specific relationship between SP2 and the TOC proteins involved in the import of photosynthesis-related preproteins – specifically, Toc33, Toc159 and Toc75-III (Jarvis and López-Juez, 2013). The

nature of this relationship remains to be determined, but it appears that SP2 has a negative regulatory effect on these TOC proteins. When *SP2* is mutated, this negative regulation is alleviated and so the suppression of chloroplast biogenesis defects in TOC mutants by *sp2* is mediated by the accumulation of the remaining TOC proteins, which enhances plastid protein import capacity. In contrast, *SP2* overexpressors enhance the negative regulatory effect, and so all three of these TOC proteins are depleted, thereby reducing protein import capacity.

Given the conservation of motifs in the β -barrel domain of SP2, and the protein's significant homology to OEP80, it is conceivable that SP2 has a function related to protein transport or assembly similar to other Omp85 proteins. The finding that SP2 could physically interact with the TOC proteins (Fig. 3.13), would support this idea. In light of the findings presented in this chapter, it appears that SP2 may have a distinctly opposite function of OEP80 (Huang et al., 2011). This could take the form of a direct regulation; for example, if OEP80 forms complexes for assembling proteins, SP2 could act to disrupt the complex due to its lack of POTRA domains, in a type of competitive inhibition. The crossing of *sp2* to *tic22-III tic22-IV* was an attempt to test this possibility, due to a speculated involvement of Tic22 in OEM protein assembly based on an apparent structural similarity to BamA cofactor, BamB. This experiment yielded no convincing level of suppression of the *tic22-III tic22-IV* phenotype by *sp2* (Fig. 3.5). However, due to a lack of evidence surrounding the role of Tic22, this may not have effectively tested this hypothesis. Recent evidence revealed an interaction of Tic22 with the POTRA domains of Toc75-III, as well as transit peptides, which is consistent with the proposal that Tic22 functions as an import chaperone (Shi and Theg, 2013b; Paila et al., 2016; O'Neil et al., 2017), however

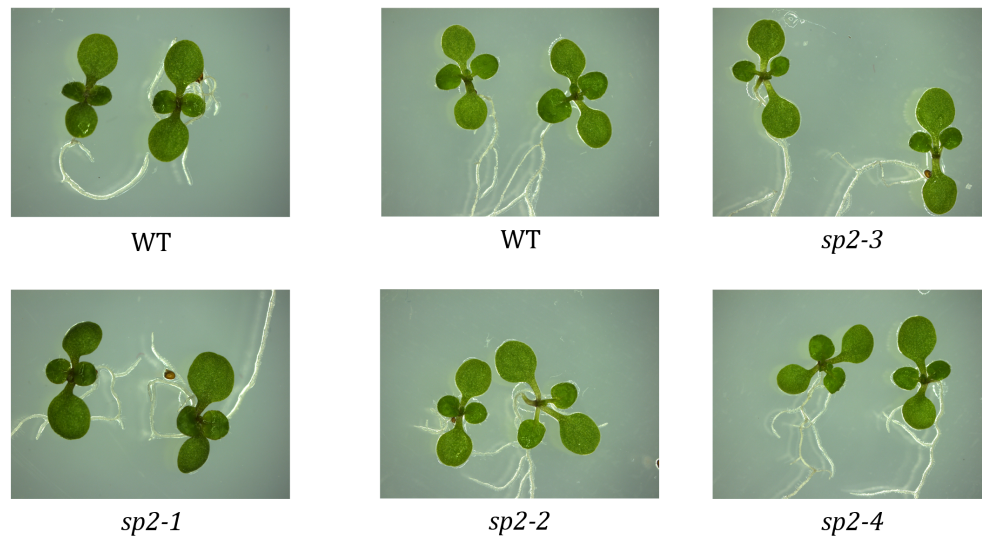
its role as a potential import chaperone does not exclude the possibility that Tic22 could chaperone unfolded β -barrel proteins or other substrates in the IMS, and so effect their assembly.

It may not be the case that SP2 has a function so close to that of OEP80, as SP2 is not simply OEP80 without POTRA domains. In Chapter 2, evidence was presented that SP2 could have neofunctionalized, and thus it may be a protein that is functionally distinct from OEP80. Considering this hypothesis, and the information presented so far, it is conceivable that SP2 acts in the assembly of, or operates with, a negative regulator of the TOC proteins. If OEP80, and to a lesser extent Toc75, act in TOC protein assembly, then SP2 could act in TOC protein removal.

The latter hypotheses call to mind the function of SP1, which effects post-translational degradation of the TOC proteins. In fact, the characterisations of SP2 so far mirror those of the SP1 very closely (Ling et al., 2012). This raises the intriguing possibility that SP1 and SP2 operate in a common pathway. This hypothesis will be explored in Chapter 4.

3.4 Figures

A



B

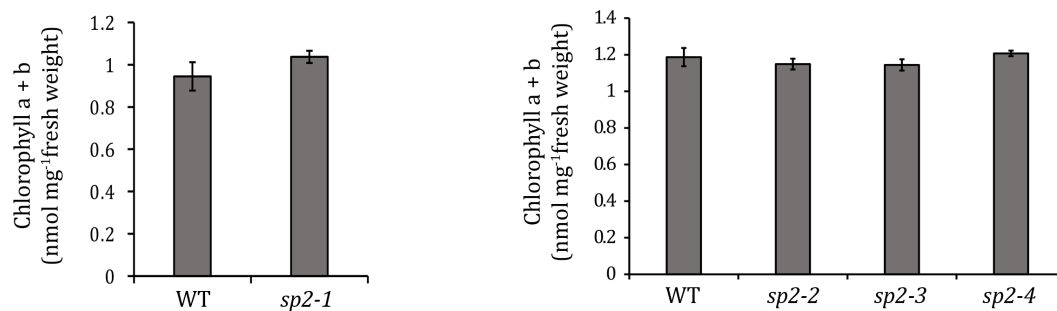


Figure 3.1. Phenotypic analysis of *sp2* mutant plants. (A) Representative images of, and (B) quantification of chlorophyll extracted with DMF from, whole 10-day-old seedlings of wild type and each of the *sp2* alleles. Note that the analysis of *sp2-1* is separated from the analysis of the other alleles, as that analysis was performed separately from the others. Error bars represent SEM (n = 4).

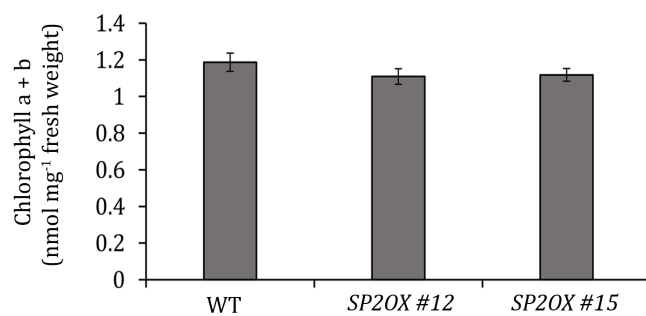
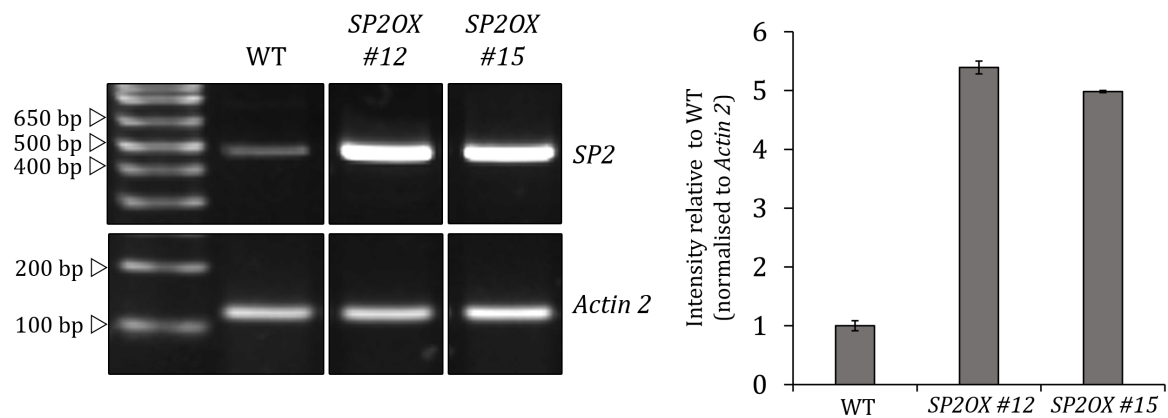
A**B****C**

Figure 3.2. Analysis of plants expressing the *SP2* CDS under the control of the cauliflower mosaic virus 35S promoter (*SP2OX*). (A) Representative images of, and (B) quantification of chlorophyll extracted with DMF from whole 10-day-old seedlings. (C) Semi-quantitative RT-PCR analysis of *SP2* in wild type and *SP2OX* plants, using *Actin 2* as a control. A representative gel image (left), alongside quantification of four biological replicates (right). Error bars in B and C represent SEM (n = 4).

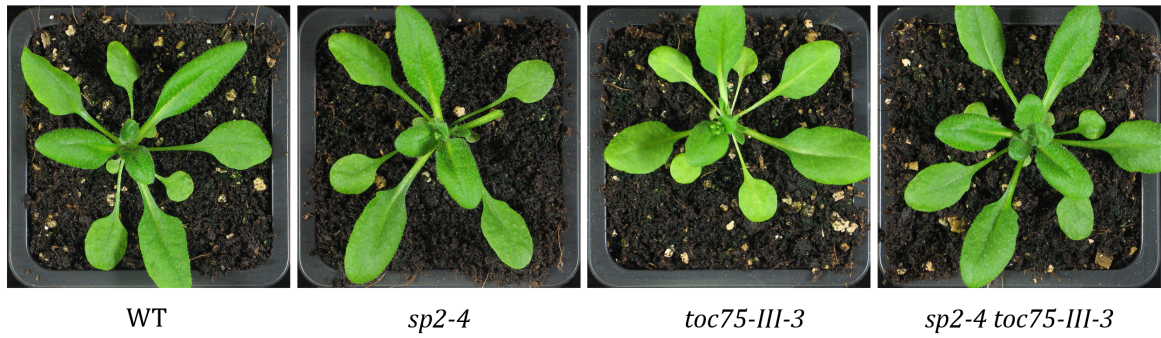
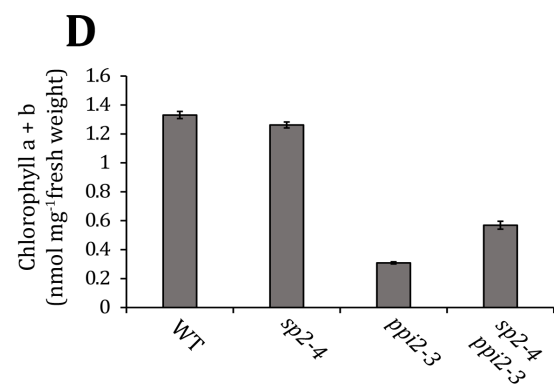
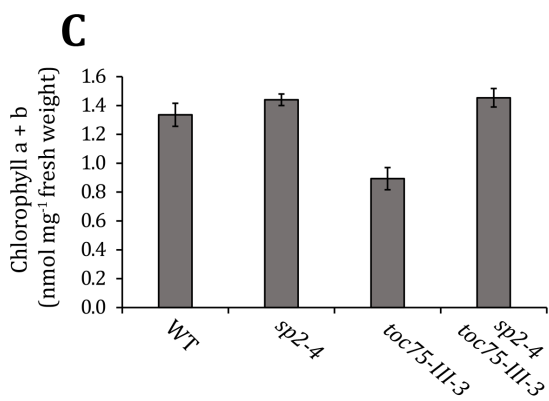
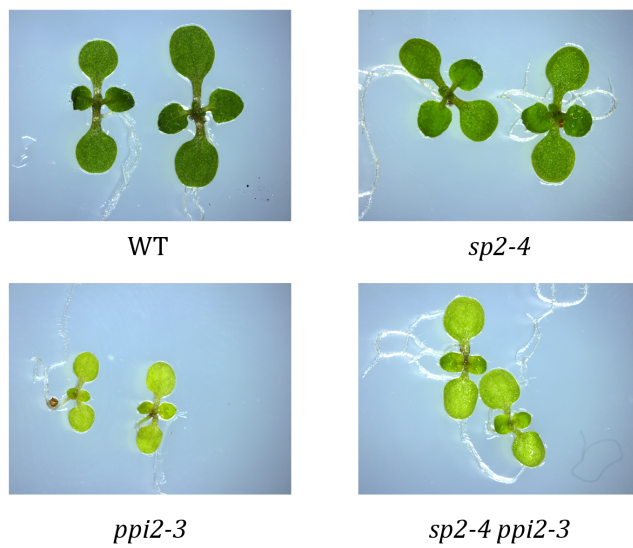
A**B**

Figure 3.3. Analysis of the effects of *sp2* on the TOC component hypomorphic mutants, *toc75-III-3* and *ppi2-3* (A and B) Representative images of 28-day-old *sp2-4 toc75-III-3* plants (A), and of 10-day-old *sp2-4 ppi2-3* plants (B), in each case alongside appropriate controls. (C and D) Quantification of chlorophyll

extracted with DMF from whole 10-day-old *sp2-4 toc75-III-3* (C) or *sp2-4 ppi2-3* (D) seedlings, in each case with appropriate controls. Error bars represent SEM (n = 4).

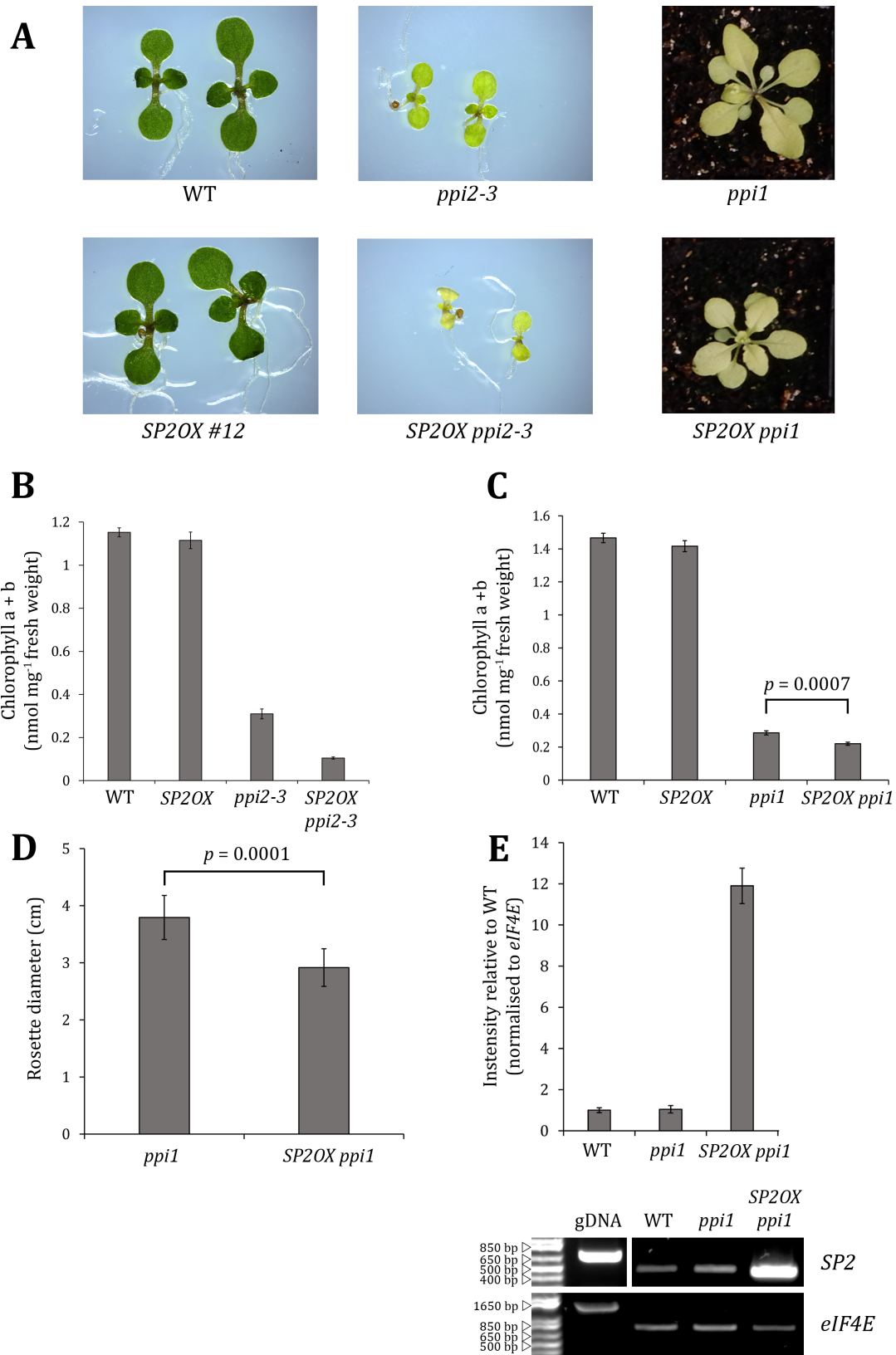


Figure 3.4. Analysis of the effects of *SP2OX* on TOC mutants, *ppi1* and *ppi2-3*. (A) Representative images of, 10-day-old *SP2OX ppi2-3* and 28-day-old *SP2OX ppi1* plants with appropriate controls. (B and C) Quantification of chlorophyll extracted with DMF from whole 10-day-old *SP2OX ppi2-3* plants (B), and

chlorophyll quantified using SPAD-502 meter measurements converted to chlorophyll concentration (Ling et al., 2011) from 28-day-old *SP2OX ppil* plants (C). Error bars represent SEM (n = 4). (D) Rosette diameter measurements from *SP2OX ppil* and *ppil* plants, as measured at the widest point. Error bars represent SD (n = 9 and 10 respectively). (E) Semi-quantitative RT-PCR analysis of *SP2* expression in *SP2OX ppil* plants normalised to *eIF4E*. Quantification representing four biological replicates (top), and representative gel image (bottom). Error bars represent SEM (n = 4). The *p* values were derived from a t-test of the means of the indicated genotypes.

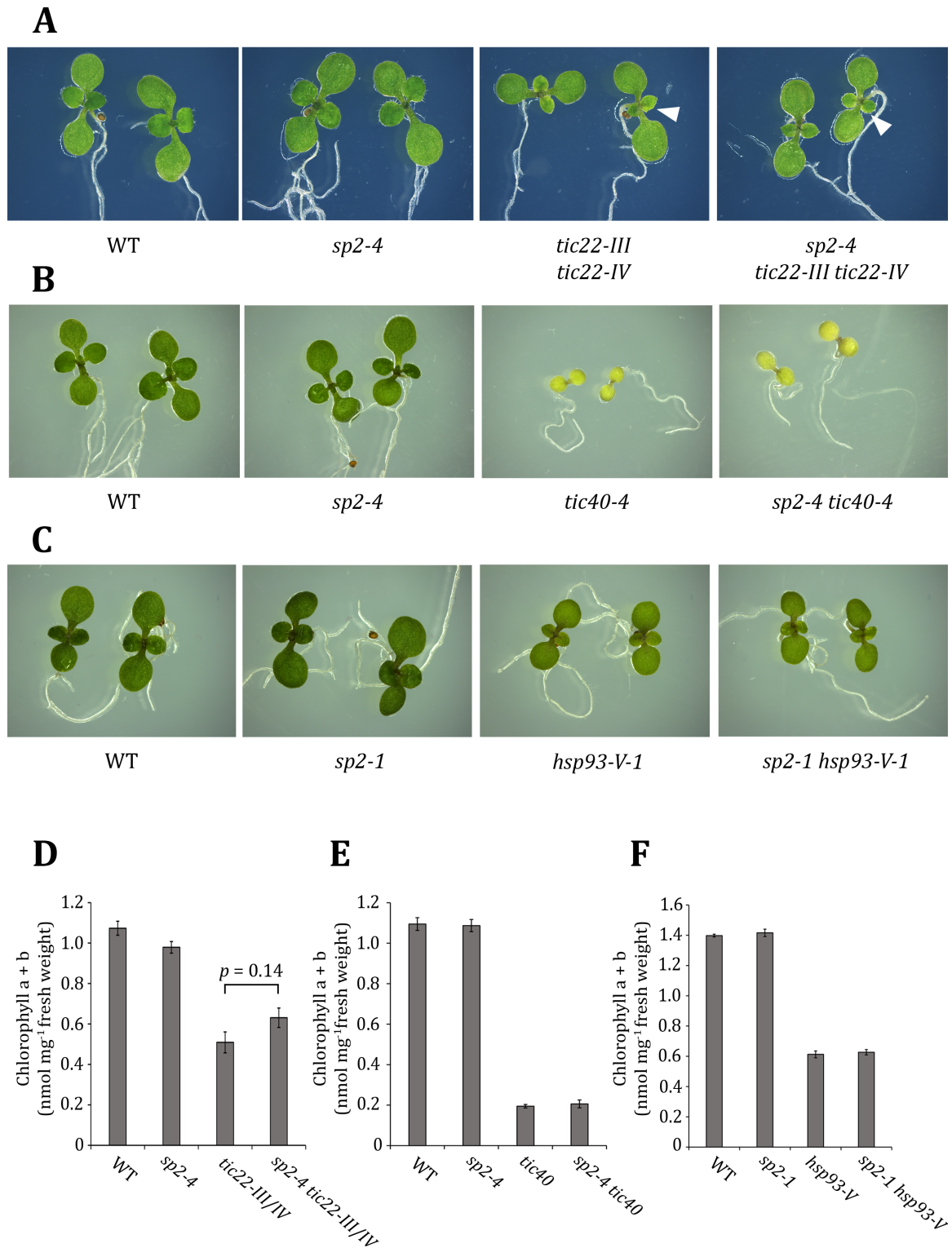


Figure 3.5. Analysis of the effects of *sp2* on mutants defective in plastid protein import. (A – C) Representative images of 10-day-old *sp2-4 tic22-III tic22-IV* seedlings (A), the white arrows indicate the white basal margins of the first true leaves, characteristic of *tic22-III tic22-IV*; *sp2-4 tic40-4* seedlings (B); and *sp2-1 hsp93-V* seedlings (C). (D – F) Quantification of chlorophyll extracted with DMF

from whole 10-day-old seedlings of, *sp2-4 tic22-III tic22-IV* (D), *sp2-4 tic40-4* (E), and *sp2-1 hsp93-V* (F). Error bars represent SEM (n = 8 (D); n = 4 (E, F)). The *p* value was derived from a t-test of the means of the indicated genotypes.

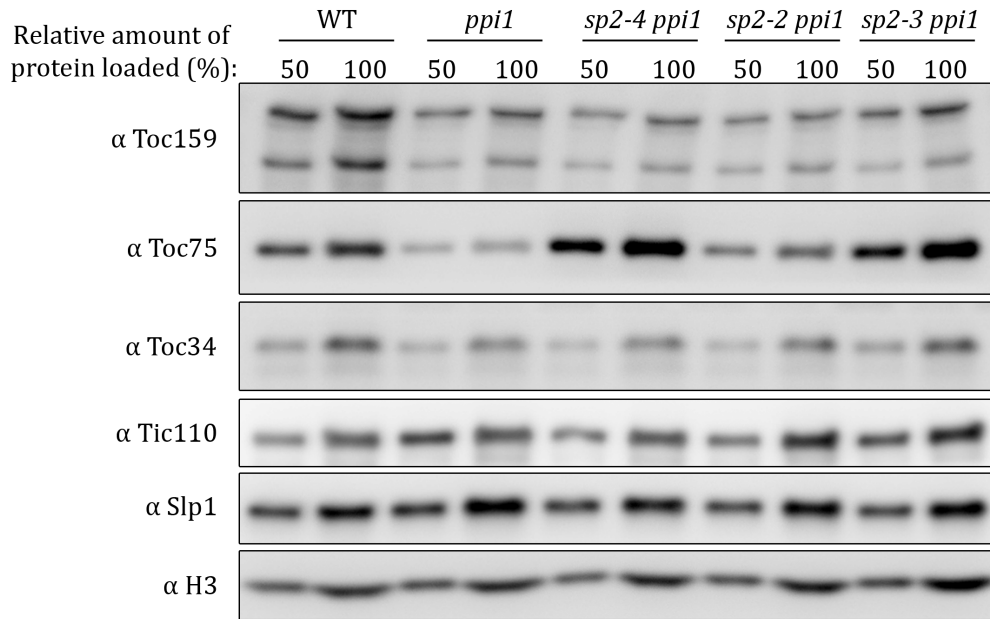
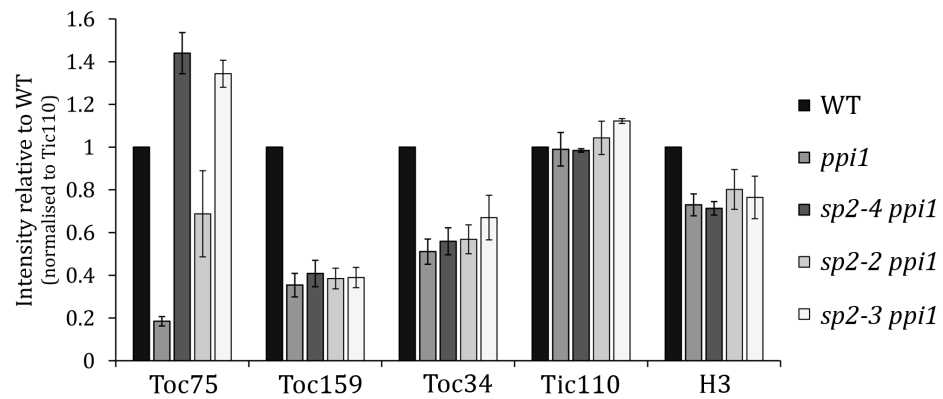
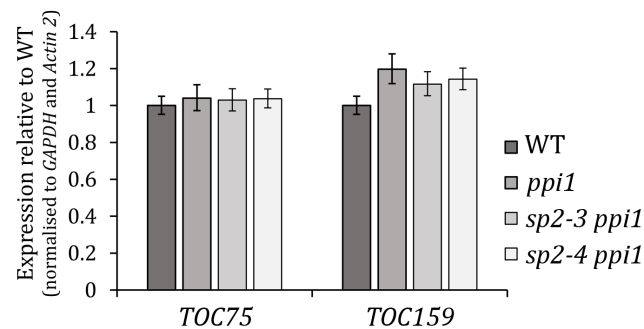
A**B****C**

Figure 3.6. Semi-quantitative western blot analysis of chloroplast protein abundance in *sp2 ppi1* mutant plants. (A) Equal amounts of total protein extracts from whole, 10-day-old seedlings were analysed using the indicated antibodies, and representative western blot images are shown. (B) Semi-quantitative analysis

of western blot images corresponding to several biological replicates. The data for each protein were normalised to corresponding data for Slp1, and are presented as a ratio to the wild type value for each respective protein. Error bars represent SEM (n = 3-4). (C) Quantitative real-time PCR (qPCR) analysis of TOC gene expression in 10-day-old wild-type, *ppi1* and *sp2 ppi1* plants, normalised to two reference genes (*Actin2* and *GAPDH*). Error bars represent SD (n = 4).

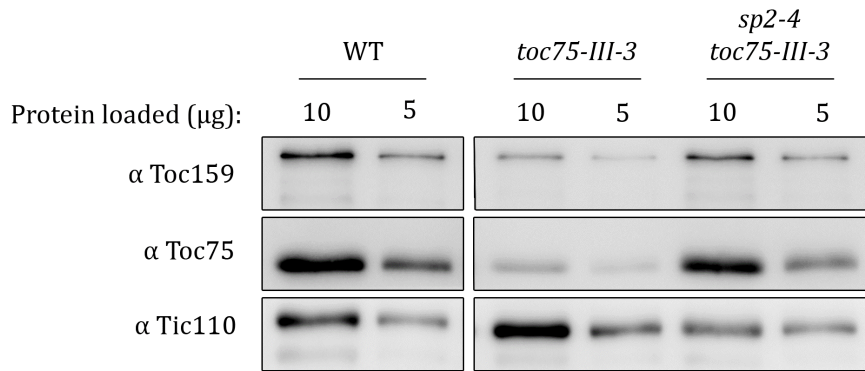
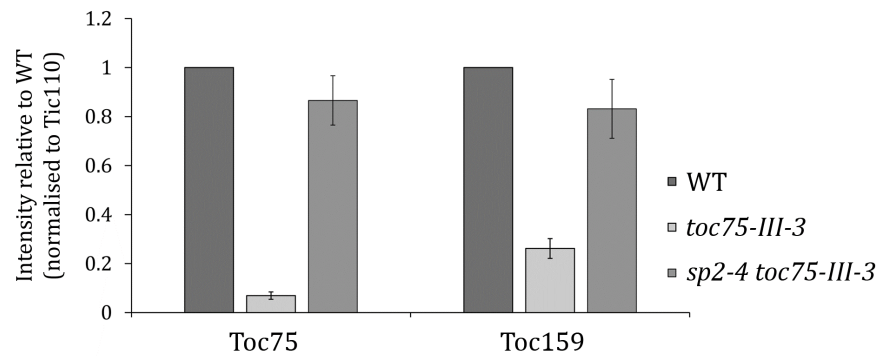
A**B**

Figure 3.7. Semi-quantitative western blot analysis of chloroplast protein abundance in *sp2 toc75-III-3* mutant plants. (A) Equal amounts of total protein extracts from whole, from 28-day-old leaf tissue were analysed using the indicated antibodies, and representative western blot images are shown. (B) Semi-quantitative analysis of western blot images corresponding to several biological replicates. The data for each protein were normalised to data corresponding to Tic110, and are presented as a ratio to the wild type value for each respective protein. Error bars represent SEM (n = 4).

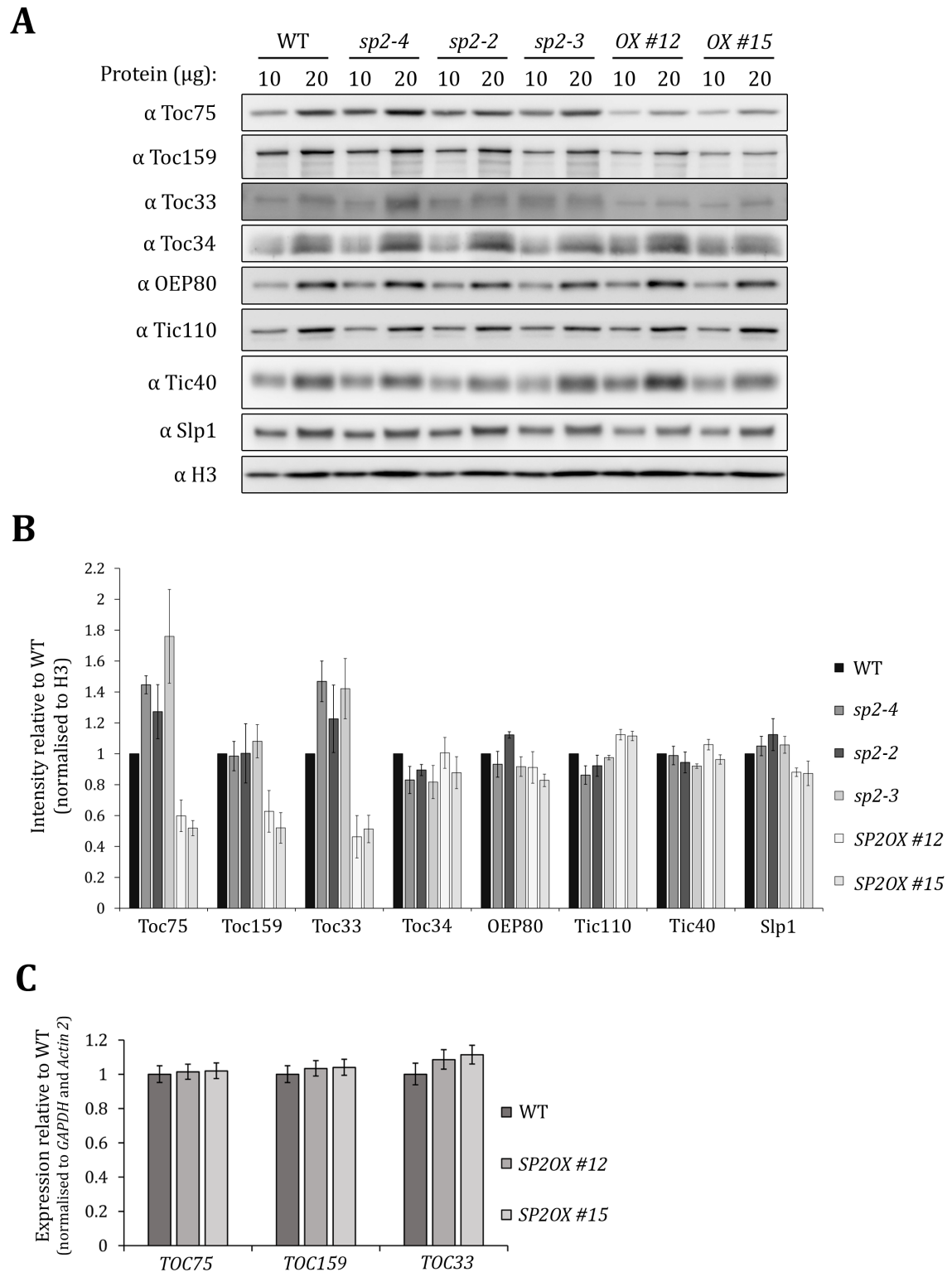


Figure 3.8. Semi-quantitative western blot analysis of chloroplast protein abundance in *sp2* mutant and *SP2OX* plants. (A) Equal amounts of total protein extracts from whole, from 28-day-old leaf tissue were analysed using the indicated antibodies, and representative western blot images are shown. *OX* =

SP2 overexpressor. (B) Semi-quantitative analysis of western blot images corresponding to several biological replicates. The data for each protein were normalised to data corresponding to H3, and are presented as a ratio to the wild type value for each respective protein. Error bars represent SEM ($n = 3 - 4$). (C) qPCR analysis of TOC gene expression in 10-day-old wild type, and overexpressor lines, #12 and #15, normalised to two reference genes (*Actin2* and *GAPDH*). Error bars represent SD ($n = 4$).

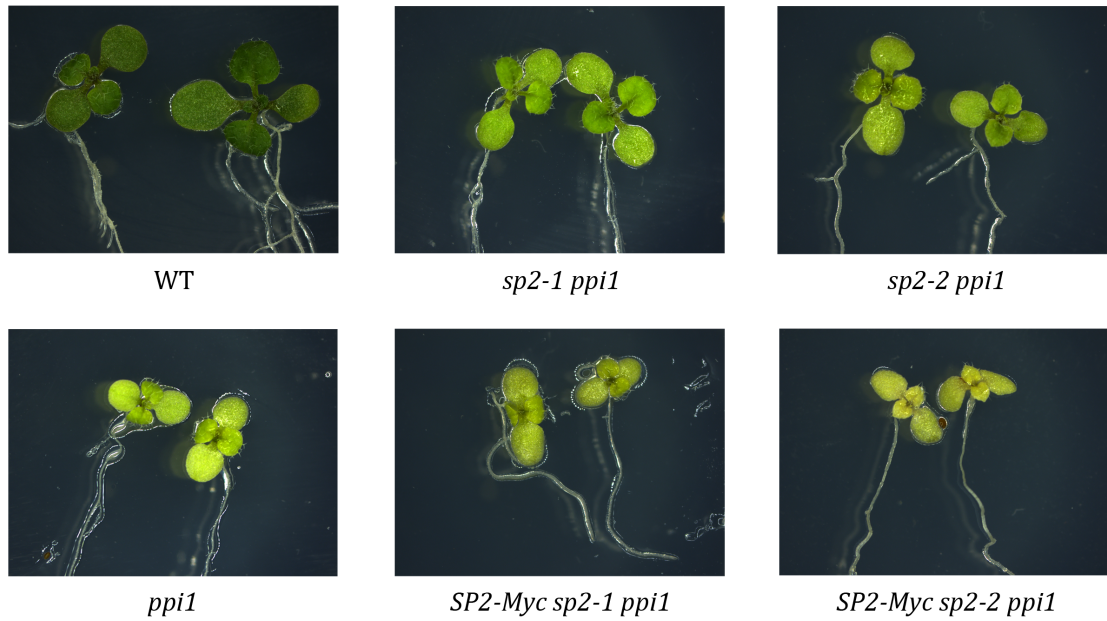
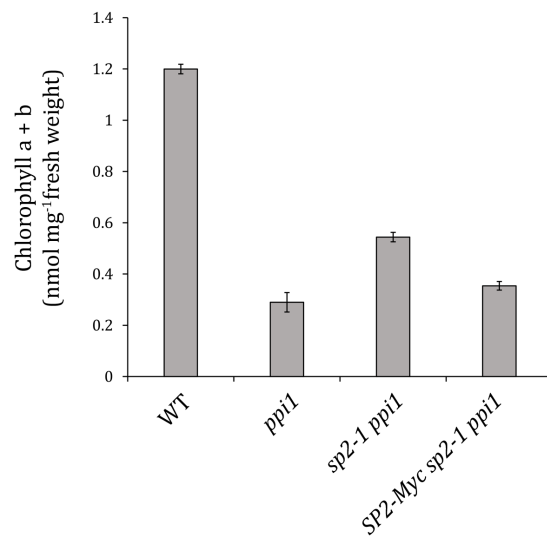
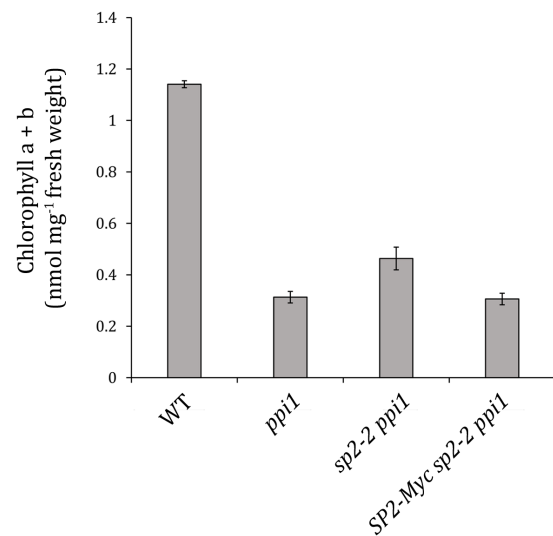
A**B****C**

Figure 3.9. Phenotypic analysis of *SP2-Myc sp2 ppi1* plants. (A) Representative images of 10-day-old plants. (B and C) Quantification of chlorophyll extracted with DMF from whole 10-day-old *SP2-Myc sp2-1 ppi1* plants (B) and *SP2-Myc sp2-2 ppi1* plants (C). Error bars represent SEM (n = 4).

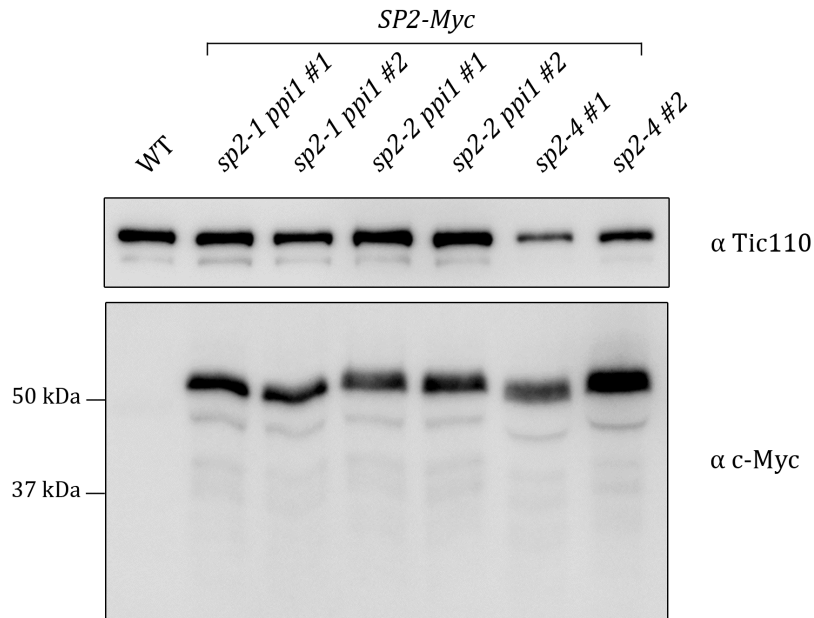


Figure 3.10. Western blot analysis of the SP2-Myc protein. Protein was extracted from whole 10-day-old seedlings of the wild type, and several different plants transformed with the *SP2-Myc* construct. These include two replicates each of *sp2-1 ppi1*, *sp2-2 ppi1*, and *sp2-1*. Equal amounts of protein from each sample were analysed by western blot, and probed with a monoclonal anti-c-Myc antibody. As a control, a western blot was also probed with an anti-Tic110 antibody.

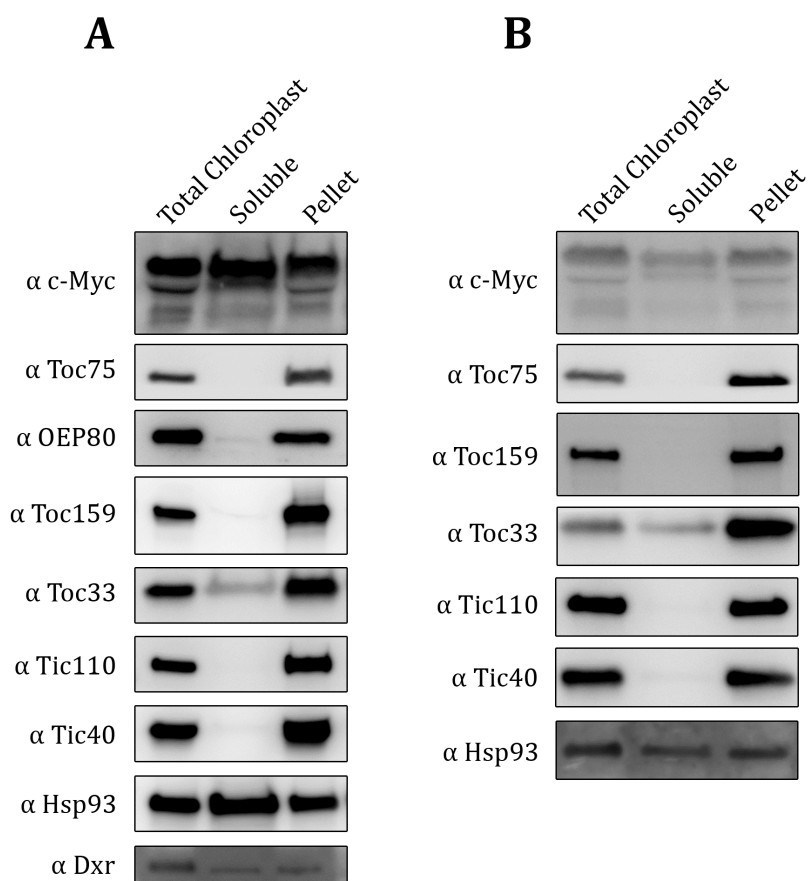


Figure 3.11. Analysis of the chloroplast membrane association of SP2-Myc by alkaline treatment of isolated *SP2-Myc sp2-4* chloroplasts. Chloroplasts isolated from *SP2-Myc sp2-4* plants were treated with Na_2CO_3 to extract soluble and peripherally membrane-associated proteins; membrane integral proteins were expected to resist such alkaline extraction. Effects of the treatment on a number of chloroplast proteins were analysed in two independent experiments: experiment 1 (A), and experiment 2 (B).

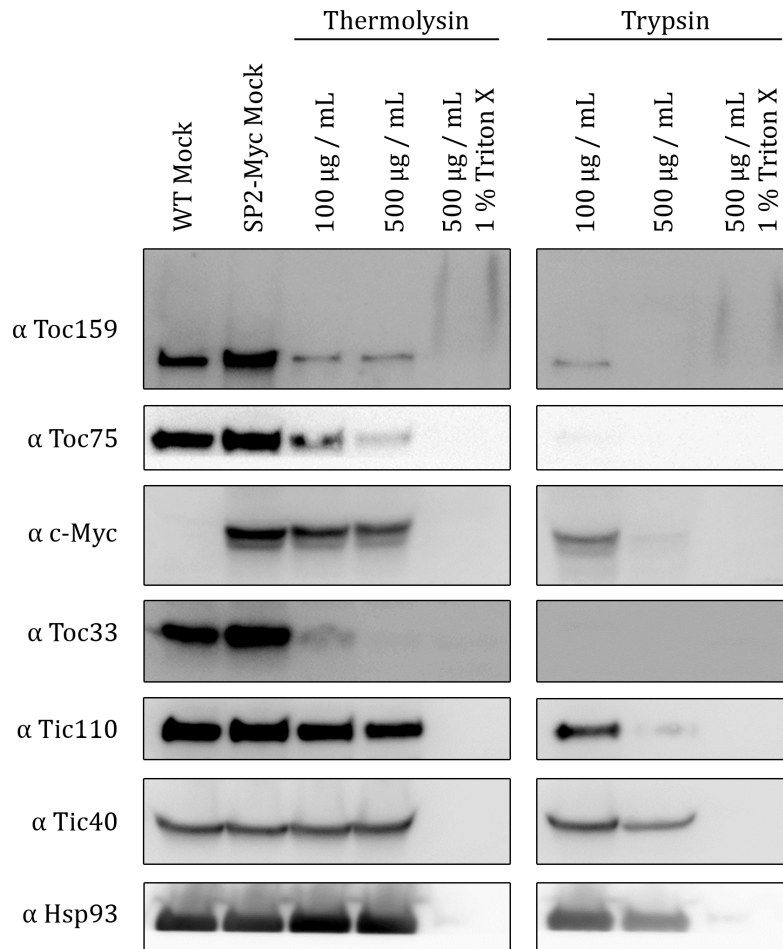


Figure 3.12. Analysis of SP2-Myc chloroplast localisation and topology by protease treatment. Chloroplasts isolated from 14-day-old *SP2-Myc sp2-4* plants were treated with thermolysin or trypsin, at different concentrations, and with the inclusion of Triton X as indicated. Effects of the treatment on a number of chloroplast proteins were analysed by immunoblotting. Proteins of with an outer-envelope localisation are expected to have limited sensitivity to thermolysin, and high sensitivity to trypsin.

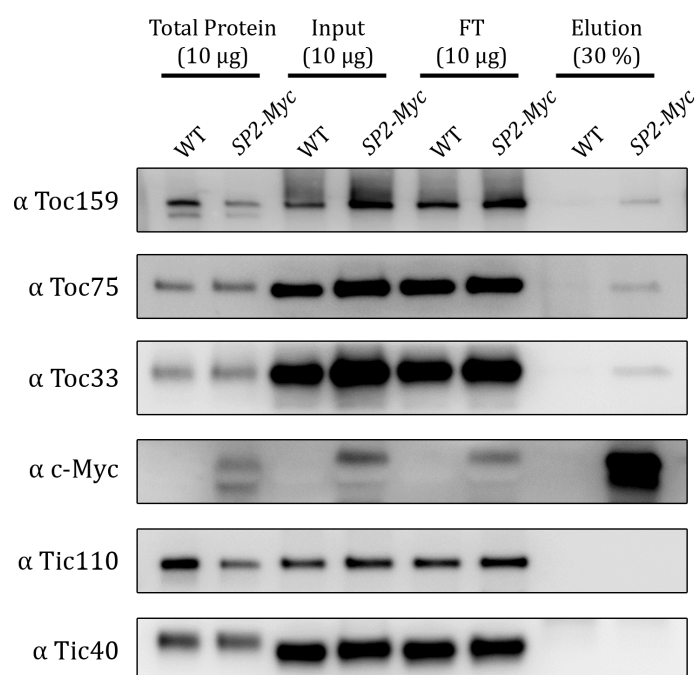


Figure 3.13. Analysis of the interaction of TOC proteins with SP2-Myc by co-immunoprecipitation. Chloroplasts were isolated from 14-day-old *SP2-Myc sp2-4* plants, lysed with hypotonic buffer, and protein was extracted for co-immunoprecipitation (co-IP) from the recovered membranes. co-IP was performed using SP2-Myc as bait. Several chloroplast proteins were analysed for possible association with SP2-Myc by immunoblotting. FT = Flow through.

Chapter 4

Results III

Interactions between SP2 and SP1

Statement of authorship

Dr Qihua Ling: Made the construct p2GW7-HA-CDC48. He also contributed significantly to understanding the function of SP1, and to developing the ChlorAD hypothesis, in particular the involvement of CDC48, discussed in this chapter.

Abstract

SP2 was discovered in a screen for suppressors of the chloroplast biogenesis mutant, *ppiL*. In Chapters 2 and 3 it was established that *SP2* has a function that is specific to the regulation of preprotein translocation across the outer-envelope membrane of the chloroplast, via negative regulation of the TOC proteins Toc33, Toc75-III and Toc159. The phenotypic and molecular effects of *SP2* mutation and overexpression were intriguingly similar to those of *SP1*. In this chapter, the relationship between the *SP1* and *SP2* functions is addressed. Genetic interaction experiments indicated that *SP1* and *SP2* operate in a common pathway, and that *SP2* acts downstream of *SP1*. As *SP1* labels TOC proteins for degradation, it is thought that a mechanism is required to remove the proteins from the membrane, as is the case for substrates of the ER-associated degradation system (ERAD). Given its channel-forming structure and the functions of its relatives, *SP2* is a strong candidate for such a role. In support of this hypothesis, co-immunoprecipitation and in vitro pulldown assays demonstrated clear interaction between the *SP1* and *SP2* proteins, and in addition *SP2* was shown to interact with ERAD ATPase, CDC48, which aids in substrate removal from the ER membrane.

4. Interactions between SP2 and SP1

4.1 Introduction

The translocon at the outer-envelope membrane of chloroplasts (TOC) is responsible for transporting into chloroplasts the cytosolically-synthesised plastid preproteins, which as mature proteins comprise >95% of the chloroplast proteome (Jarvis and López-Juez, 2013). Multiple lines of evidence point to a need to regulate the TOC complex, which can be composed of various different sets of components (Shi and Theg, 2013b; Demarsy et al., 2014; Ling and Jarvis, 2015a).

To probe for undiscovered factors influencing the function and composition of the TOC and its constituent components, a mutagenesis screen was conducted for second-site suppressors of the Toc33 mutant, *ppi1*. This screen yielded two suppressor loci, *suppressor of ppi1 locus 1 (sp1)*, described by Ling et al. (2012), and *suppressor of ppi1 locus 2 (sp2)*, described in Chapters 2 and 3 here.

In Chapter 2 it was found that SP2 is an Omp85 β -barrel protein, closely related to OEP80, and in Chapter 3 it was discovered that the function of SP2 most likely acts in the opposite capacity to that of OEP80 which is thought to assemble plastid OMPs. Despite a high degree of homology between SP2 and OEP80, and due to a lack of information on OEP80 and SP2 functions, it is presently uncertain if these two proteins are functionally related.

SP2 was shown to influence the abundance of the TOC proteins linked to the import of highly-abundant photosynthesis-related preproteins, Toc33, Toc75

and Toc159. This evidence was indicative of a strong negative regulatory influence of SP2 on these TOC proteins, and that this likely happens post-translationally. The mechanism underlying the suppression of the TOC mutants by *sp2*, therefore, is that the remaining TOC proteins are able to accumulate to a higher abundance and so can partially compensate for the protein import defect.

As OEP80 was proposed to have a function in the assembly of outer-envelope proteins, similar to its relatives in bacteria (Huang et al., 2011; Töpel et al., 2012), one hypothesis is that SP2 could exhibit a regulatory influence on OEP80. Also, given the homology between SP2 and OEP80, SP2 is likely to retain a channel function. In fact, the SP2 channel properties have already been studied by electrophysiology, which showed that the conductance and pore-size were similar to those of OEP80 and Toc75 (Hsueh et al., 2017). Therefore, alternative hypotheses are that SP2 functions in the extraction of membrane proteins including those of the TOC complex, that it assembles membrane proteins which exhibit negative regulatory function on the TOC, or that it exhibits some sort of competitive activity against the TOC.

Interestingly, the effects of SP2 on chloroplast biogenesis and to plastid proteins recorded so far are very similar to those of SP1 (Ling et al., 2012). SP1 is a chloroplast outer-envelope membrane-localised E3 ubiquitin ligase, and a component of the ubiquitin-proteasome system. An E3 ubiquitin ligase ligates ubiquitin to target proteins to label them for degradation by the cytosolic 26S-proteasome (Vierstra, 2009). In the case of SP1, its targets have been shown to be the TOC proteins (Ling et al., 2012; Ling and Jarvis, 2015b). It was found that SP1 activity was important for influencing plastid transitions, and for abiotic stress tolerance (Ling et al., 2012; Ling and Jarvis, 2015b). This presumably allows for

TOC rearrangement or modulation, and there is evidence for this, as the relevant studies demonstrated changes in TOC abundance at the protein level under these conditions.

In the case of plastid transitions, different plastid types must import different sets of preproteins, and different TOC components are associated with different preprotein clients; thus, the rearrangement of the TOC apparatus to facilitate the import of these preproteins is required (Jarvis and López-Juez, 2013; Shi and Theg, 2013b; Demarsy et al., 2014). In the case of abiotic stress, the chloroplast can be subject to oxidative stress caused by an imbalance between the capacity for light harvesting and electron transport, and the limited availability of CO₂ as an electron acceptor, due to closure of stomata (Sharma et al., 2012). These stress factors can be reduced by preventing maintenance and biogenesis of the photosystems and oxygen evolving complexes, which can be achieved in part by inducing TOC degradation, to prevent import of new photosynthetic components translated in the cytosol (Hua and Vierstra, 2016; Ling and Jarvis, 2016).

Given the similar functional characteristics of *SP1* and *SP2*, it was important to test the interaction between these two genes/proteins, as it was conceivable that they may act in a common pathway. Presence or absence of such an interaction would suggest several hypotheses:

- (1) *SP2* acts upstream of *SP1*, and forms a highly-specific protein assembly pathway for negative regulators of the TOC, including *SP1*, and possibly its chloroplast-localised homologues, *SP1-like 1* (*SPL1*) and *SPL2*.
- (2) *SP2* acts downstream of *SP1*, and operates in the removal of ubiquitin-targeted TOC proteins.

(3) Alternatively, if there is no interaction between the components, then SP1 and SP2 may act in independent pathways of TOC regulation.

The second hypothesis would mean that SP2 fulfils a function similar to that of Hrd1 in the ERAD pathway (Baldrige and Rapoport, 2016); see Section 1.3 for explanation of ERAD. When membrane proteins are targeted for degradation, there is a requirement for energy and a pathway to exit the membrane (Avci and Lemberg, 2015). SP2 could facilitate this process, acting as a retrotranslocon.

In fact, a cytosolic AAA+ ATPase called CDC48 provides the motive force for substrate protein extraction in ERAD (via Hrd1, which fulfils the role of both E3 ligase and retrotranslocon; Lemus and Goder, 2014; Baldrige and Rapoport, 2016), and in the recently-described mitochondrial-associated degradation system (MAD; Section 1.4; Wu et al., 2016). Data from the Jarvis lab also reveal an involvement of CDC48 in the SP1- and proteasome-dependent degradation of TOC proteins (unpublished). Thus, SP2 could fulfil an important role, as retrotranslocon, in the emerging ChlorAD. The requirement for a translocon is due to SP1 having only two transmembrane spans, and so, unlike the ERAD E3 ligase Hrd1 which has six transmembrane spans, it probably could not itself form a retrotranslocation channel (Baldrige and Rapoport, 2016); that is, unless SP1 forms a multimeric complex to create the channel.

Alternatively, SP2 could simply have a separate role not directly related to the function of SP1, in the regulation of TOC proteins. This chapter will explore further the function of SP2, by testing for its interaction with SP1 genetically and physically at the protein level, as well as with CDC48.

4.2 Results

4.2.1 Test of the similarity between *sp1* and *sp2* phenotypes

It has been stated that the so-far recorded effects of the mutation or overexpression of *SP2* are highly similar to those observed for *SP1* (Ling et al., 2012). Therefore to test this idea, more specific phenotypic similarities were sought. It was previously reported that *sp1* has a sensitivity to salinity (Ling and Jarvis, 2015b), and so this was tested in comparison to *sp2*. Wild type, *sp1* and *sp2* plants were placed on media containing 150 mM NaCl, and this caused clear developmental retardation of *sp2* to a similar degree as *sp1*, in comparison to wild type, as measured by chlorophyll content (Fig. 4.1). As the germination rates on the high salinity media for the tested seed were similar for all three genotypes (Fig. 4.1), it is unlikely that the effects were related to seed quality or germination, but with plant development post-germination. This observation taken with those in Chapters 2 and 3, therefore gives yet further support for a possible functional relationship between *SP1* and *SP2*, and that they could act in the same pathway.

4.2.2 Analyses of genetic interactions between *SP1* and *SP2*

4.2.2.1 Together, *sp1* and *sp2* do not exhibit phenotype additivity

As *SP1* and *SP2* share several characteristics, as discussed earlier, it was conceivable that they may act in the same pathway. In order to test this possibility, *sp1-1 ppil* was crossed with *sp2-1 ppil*, to generate the *sp1-1 sp2-1 ppil* triple mutant.

Having identified the triple mutant, its phenotype was characterised to assess for possible additivity or epistasis. First, it should be noted that there was

no apparent difference in the chlorophyll content between *sp1-1 ppil* and *sp2 ppil*, despite subtle differences in visual appearance (Fig. 4.2). Significantly, when brought together in the triple mutant, *sp1* and *sp2* did not cause enhanced suppression of *ppil*, nor was there an increased severity of the phenotype, and this is demonstrated objectively in the chlorophyll content which was similar between *sp2 ppil* and *sp1 sp2 ppil*, and between *sp1 ppil* and *sp1 sp2 ppil* (Fig. 4.2B).

One possible interpretation of these data is that, after a certain point, there are diminishing returns in terms of *ppil* phenotype suppression gained from an increase in the abundance of TOC proteins. To address this possibility, the abundance of Toc75-III and Toc159 in the triple mutant was analysed by western blotting (Fig. 4.2). The results revealed that the abundance of both TOC proteins is similar between *sp1 ppil*, *sp2 ppil* and *sp1 sp2 ppil*. The fact that the control proteins, Tic110 and nuclear H3, were present at similar levels in all genotypes tested provided a clear indication of proper sample normalization (Fig. 4.2).

Thus, the ‘diminishing returns’ interpretation was ruled out. Nonetheless, phenotype additivity could be masked in the triple mutant, for example resolution of the effects to the abundance of TOC proteins could be limited by the total amount of TOC protein synthesised by the point in time during the development of the plant at which the protein samples were taken.

The observed lack of phenotypic additivity is also consistent with either *SP1* or *SP2* being epistatic over the other, but as the *sp1 ppil* and *sp2 ppil* both have similar chlorophyll contents (Fig. 4.2B) and TOC protein abundance (Fig. 4.3), it is not possible to resolve which. Thus, further steps were taken to elucidate the nature of the interaction between *SP1* and *SP2*.

4.2.2.2 The phenotype of *SPIOX ppil* is suppressed by *sp2*

The next approach was to test whether overexpression of either *SP1* or *SP2*, could compensate for mutation of the other. Or in other words, could *sp1* or *sp2*, still suppress *ppil* if *ppil* was overexpressing the other component. If *SP1* and *SP2* are dependent upon each other for their functions in TOC protein regulation, then overexpression of one component would not be predicted to compensate for the absence of the other. For example, if *SP2* is required for the removal of TOC proteins for degradation, when *ppil* overexpresses *SP1* (which ubiquitinates the TOC proteins to signal their proteasomal degradation, and increases the severity of *ppil*), then *sp2* should still be able suppress the severe phenotype of *SPIOX ppil*, as it does to *ppil* plants; this would be because, despite the ubiquitination of TOC proteins, without *SP2* there would be no pathway for TOC protein removal. An alternative scenario can be envisaged where *SP2* is required upstream of *SP1* activity, for example in *SP1* assembly; in this case, *SP1* would be unable to assemble and target proteins for degradation in the first place. However, it is worth noting that the results could be complicated if it were the case that *SP2* is not the only protein-extraction mechanism, or if *SP1* can spontaneously assemble but at a lower rate than with *SP2*'s assistance.

To test these hypotheses, *SPIOX* was crossed to *sp2 ppil*, and triple homozygotes were identified. Whereas it was possible for the *SPIOX* transgene to accentuate the *ppil* phenotype (Fig. 4.4; Ling et al., 2012), there was no statistically significant difference in chlorophyll content between *sp2 ppil* and *SPIOX sp2 ppil* plants (Fig. 4.4B). Therefore as predicted by the hypothesis, *sp2* was able to suppress *SPIOX ppil* (Fig. 4.4), and so *SPIOX* was unable to

compensate for lack of SP2. This conclusion would support the notion that SP2 is epistatic to SP1, i.e., that SP2 could act downstream of SP1.

However, it should be noted that the expression of *SP1* in *SP1OX ppil* plants was >7-times that in wild type, and 11-times that in *ppil*, as determined by RT-PCR, whereas the expression of *SP1* in *SP1OX sp2 ppil* plants is only 2.5-times that in wild type, and 4-times that in *ppil* (Fig. 4.4). Thus, it could be argued that these plants did not produce enough SP1 to cause reversion of the suppression effect. That said, the expression of *SP1* in *SP1OX sp2 ppil* plants was comparable to that seen in the *SP1OX ppil* plants which did exhibit phenotype accentuation, as reported previously by Ling et al. (2012).

4.2.2.3 The phenotype of *SP2OX ppil* is not suppressed by *sp1*

In the reciprocal experiment, *SP2OX* was crossed to *sp1 ppil*, and again triple homozygotes were identified and analysed. This interaction was slightly more complex. The *SP2OX* transgene only subtly accentuates the *ppil* phenotype (Fig. 4.5C, D), despite the fact that *SP2* was highly overexpressed (Fig. 4.5E). Interestingly, however, *sp1* was unable to suppress to *SP2OX ppil* as it does *ppil*, and the chlorophyll content was significantly reduced in *SP2OX sp1 ppil*, to 43% that of *sp1 ppil* (Fig. 4.5B), and so *SP2OX* could compensate for absence of SP1. This was not specifically predicted by the hypothesis, but it is still consistent with the hypothesis. There is a plausible explanation for the inability of *sp1* to suppress *SP2OX ppil*. It could be that the SP1 homologues, SPL1 and SPL2 (or other E3 ligases), are able to compensate for the lack of SP1, but for some reason SP2 is rate limiting (or of low abundance) in the *sp1 ppil* mutant. This interpretation also places SP2 downstream of SP1, as predicted by the ChlorAD hypothesis.

4.2.2.4 SP1 and SP2 overexpression yields a synergistic phenotype

To test the SP1-SP2 interaction further, *SP1OX SP2OX* double-overexpressor plants were generated by crossing. The parent *SP1OX* and *SP2OX* plants both exhibited a normal, wild-type visible phenotype, and yet both had clearly decreased levels of TOC proteins (Fig. 4.6). However, when brought together in *SP1OX SP2OX* plants a severe phenotype resulted (Fig. 4.6). The plants appeared small and pale-yellow, reminiscent of *ppi1*, and contained 35% of the normal chlorophyll content, as compared to wild type (Fig. 4.6B). The visible phenotype of *SP1OX SP2OX* is clearly reflected by the severe depletion of the TOC proteins, which are significantly reduced in abundance in comparison to *SP1OX* or *SP2OX* alone, whereas Tic110 appears unaffected. This result can be interpreted as gene synergy, both with respect to the visible phenotype and the effects to protein abundance.

4.2.3 Analyses of the protein-protein interactions of SP2

The SP1-SP2 genetic interactions studies left open different possible interpretations, but certainly did not contradict the possibility of a direct physical interaction between the components. Thus, it was necessary to analyse directly whether there is an interaction between SP1 and SP2. To this end, co-immunoprecipitation and in vitro pulldown studies were pursued. Due to the lack of an SP2 antibody, an epitope-tagged version of SP2 was used in the co-immunoprecipitation studies, whereas for the in vitro pulldown experiment, the SP2 cDNA was translated in vitro, with radiolabelled methionine.

4.2.3.1 Co-immunoprecipitation analysis of SP2

As *SP2* had been shown to interact genetically with mutations affecting the TOC genes (Chapter 2), and with *SP1* (Figs. 4.4, 4.5, 4.6), it was important to investigate the nature of these interactions at the protein level. Because the *SP2-Myc* construct could complement *sp2 ppi1* mutants, and because localisation studies indicated a chloroplast localisation for SP2-Myc (Figs. 3.9, 3.11, 3.12), co-immunoprecipitation studies with these plants were pursued.

Using beads conjugated with anti-c-Myc antibodies, protein samples from isolated chloroplasts of *SP2-Myc sp2-4* plants was analysed to assess for TOC protein association with the SP2-Myc bait. In these experiment, SP2-Myc specifically pulled down the TOC proteins Toc33, Toc75-III and Toc159, none of which was pulled down from equivalent wild type samples (Fig. 3.10). The SP2-Myc bait did not pulldown chloroplast IEM proteins Tic110 or Tic40 (Fig. 3.10). These data indicate a specific physical interaction with the TOC proteins, consistent with the genetic interaction data described in Chapter 2, and imply that SP2 exerts negative regulation over these TOC proteins.

Next, the possibility of a physical interaction between SP2 and SP1 was explored. In addition, a possible interaction between SP2 and CDC48 was tested, motivated by the evidence that CDC48 acts in the proteasomal degradation of TOC proteins (unpublished), by analogy with ERAD and MAD. To this end, a construct carrying *SP1* fused C-terminally with the *HA* epitope tag under the control of the cauliflower mosaic virus 35S promoter (Ling et al., 2012) was used for transient expression in *Arabidopsis mesophyll* protoplasts (TEAMP) experiments (Wu et al., 2009), the protoplasts for which were obtained from *SP2-Myc sp2-4* plants. Using c-Myc antibody-conjugated beads, SP1-HA was

specifically co-immunoprecipitated with SP2-Myc (Fig. 4.7A). In control experiments, SP1-HA was not pulled down by the beads themselves from the wild type protoplast protein, lacking SP2-Myc (Fig. 4.7A), nor was YFP-HA (Ling et al., 2012) pulled down by SP2-Myc (Fig. 4.7A, B).

Using the same TEAMP system, similar results were obtained when an HA-CDC48 fusion was transiently expressed instead of SP1-HA. The CDC48 protein with an N-terminal HA epitope tag was specifically co-immunoprecipitated with the SP2-Myc bait (Fig. 4.7B). Thus, SP2 was shown to physically interact with all of the thus-far identified components of the ChlorAD system, as well as some important substrate proteins.

4.2.3.2 In vitro pulldown of SP2 by SP1

To further investigate the links between SP1 and SP2, the interaction of these two proteins was studied in vitro. For this work, SP1 was fused to an N-terminal GST tag and was expressed in, and purified from, bacteria. In the relevant construct, called *GST-SP1flex*, the two transmembrane spans of SP1 were deleted: the transmembrane domain at the N-terminus was removed, and the transmembrane domain between the IMS domain and C-terminal RING domain was replaced with a flexible linker, with the sequence 2xGGGGS (Ling et al., 2012). This was to aid in solubility, and to minimise the incidence of non-specific interactions due to the presence of these hydrophobic domains. In parallel, the native SP2 protein was translated in vitro, and labelled with radiolabelled [³⁵S]-methionine, using a rabbit reticulocyte lysate system.

Using these reagents in pulldown experiments in vitro, GST-SP1flex bound to glutathione agarose beads specifically bound radiolabelled SP2 (Fig. 4.8);

specificity was demonstrated by the fact that GST alone bound to glutathione agarose beads did not pull down radiolabelled SP2 (Fig. 4.8). Unfortunately, as SP2 is exclusively a β -barrel protein, there were no prominent hydrophilic domains which could be analysed separately in further experiments to define the determinants of the interaction. While it is possible that the interaction of SP2 with GST-SP1flex in this assay could have been through non-specific hydrophobic interactions between the two proteins, this seems unlikely in view of the strength and specificity of the interaction (in this regard, it is noteworthy that the GST control was strongly overloaded relative to GST-SP1flex). Further experimentation with other GST fusion proteins to test whether SP2 interacts with the bait proteins non-specifically will help to resolve this.

In conclusion, physical interactions of SP2 were observed with SP1 *in vitro* and *in vivo*, and with the TOC protein extraction-related component CDC48.

4.3 Discussion

The discovery of the similarities between the *SP2* and *SP1* mutant and overexpressor phenotypes, and the genetic interactions of *SP2* with genes encoding Toc33, Toc75-III and Toc159, was interesting. It meant that there was a possible interaction between SP1 and SP2, and that the two proteins may operate in a common pathway to facilitate the degradation of the TOC proteins. In particular, SP2 could fulfil the role of a hypothesised channel required to extract chloroplast membrane proteins labelled for degradation.

In this chapter, the possibility of an interaction between SP2 and SP1 was investigated directly. This was done by testing more specifically the similarity in phenotypes of *sp1* and *sp2*; *sp2* like *sp1* displayed a sensitivity to high salinity (Fig 4.1), which provides a further indication that the *sp2* mutation affects the same pathway as *sp1*, and for the same biological purpose. Then their genetic interactions were assessed, and finally the physical protein-protein interactions between SP2 and SP1 was tested. Subsequently, the potential of SP2 to contribute to a hypothesised pathway of chloroplast envelope protein degradation, dubbed chloroplast associated degradation (ChlorAD), analogous to the ERAD pathway, has begun to be explored and tested.

4.3.1 SP2 and SP1 function in a common pathway

Phenotypic assessments of SP2 closely mimicked SP1, and was specifically tested in relation to salt-stress (Fig 4.1), which was again consistent with their participation in a common pathway. Further to this, the genetic interaction data revealed a lack of additivity or synergy when the *sp1* and *sp2* mutations were combined in the *ppi1* background to make a triple mutant (Fig. 4.2). This was

reflected in the levels of TOC proteins, which did not change in these plants relative to the *sp_n ppil* double mutants (Fig. 4.3). These data were supportive of the notion that *SP1* and *SP2* do not operate in separate pathways, but in the same pathway to affect TOC protein abundance.

If *SP1* and *SP2* do act in the same pathway, then one might expect that *sp2* / *sp1* should be able to suppress and *SP1OX ppil* / *SP2OX ppil*, respectively. As expected, *sp2* was able to suppress *SP1OX ppil* (Fig 4.4); the phenotype and chlorophyll content of *SP1OX sp2 ppil* plants appeared the same as *sp2 ppil*, indicating that *SP2* may be epistatic to *SP1* (Fig. 4.4). In the converse experiment, *sp1* did not suppress *SP2OX ppil* plants, such that the visible phenotype and chlorophyll content of the *SP2OX sp1 ppil* plants were much paler than *sp1 ppil*, and appeared similar to *ppil* (Fig. 4.5). If *SP1* and *SP2* do act together, this would mean that *SP2* could act downstream of *SP1*, but this result is puzzling, as this indicates that *SP2* can act independently of *SP1*. However, it may be explained by the presence of *SP1*-like proteins, *SPL1* and *SPL2*, in the chloroplast envelope membrane (Ling et al., 2012): the abundance of these proteins may be stabilised, or their activity could be enhanced, by the presence of *SP2*. Additionally, it is worth noting that if *SP2* acts downstream of *SP1*, that in the *SP1OX sp2 ppil* plants, the TOC proteins would be expected to be ubiquitinated, yet given that the *SP1OX ppil* phenotype can be suppressed, the ubiquitin-modification of the TOC proteins is not inhibiting their protein transport function.

From these experiments, one may conclude that *SP2* is epistatic to *SP1*, and so may act downstream of *SP1*. The observation that *SP2OX* cannot accentuate the phenotype of *ppil*, in the same way that *SP1OX* can, is consistent with this idea (Fig. 4.5). In this case, the activity of *SP2* in removing TOC proteins

would be dependent upon the upstream activity of the E3 ligases ubiquitinating the substrates. It is likely that the E3 ligases in *ppi1* are already over active; the ability of *sp1* and *sp2* to cause a 5-6-fold increase in Toc75-III in *ppi1*, but only a maximum two-fold increase in Toc75-III in the wild-type background (Figs. 3.6, 3.8) is evidence of this. This is possibly because the lack of Toc33 affects the ability to form the heteromeric TOC complex from the remaining components, and may render the TOC complexes which do form unstable, such that TOC components become more accessible to SP1 or to other chloroplast membrane E3 ligases. Therefore, the *SP2OX ppi1* negative result can be explained by supposing that there is already sufficient SP2 in *ppi1* to facilitate TOC protein degradation, so that the increase in abundance of SP2 does not enhance the already over-active degradation activity of the E3 ligases. However, seemingly in contradiction to the previous interpretation, overexpression of *SP2* in *ppi2-3* (a Toc159 mutant) did result in phenotype accentuation. But this may be explained by the fact that Toc159 is still present, albeit with a portion of its M-domain truncated (as described in Chapter 3; Woodson et al., 2015), and is at the heart of the TOC complex (Shi and Theg, 2013b), and so may not cause the same TOC complex instability, and increased E3 ligase accessibility, as *ppi1*.

Consistent with the interpretation of *SP1* and *SP2* cooperation was the severe phenotype caused by the overexpression of both genes together, in stark contrast to the wild-type phenotype observed when they were overexpressed individually (Fig. 4.6). In individual overexpressors, the rate of TOC protein degradation may be limited by low abundance of other components (which could even be actively reduced in the overexpressor, e.g. by negative feedback regulation of the component in question), and so the TOC components are not

diminished so much as to produce chlorotic or growth retardation phenotypes; however, in the double overexpressor, such limitations may be alleviated. This is expected from the overexpression of multiple components in the same metabolic pathway; both increased metabolic flux and enhanced phenotypes are common outcomes (Veen et al., 2003; Yang et al., 2009).

The notion of direct functional co-operation between SP1 and SP2 was further supported by the clear results obtained from protein-protein interaction studies. In co-IP experiments using SP2-Myc as bait, a strong interaction was revealed between SP2-Myc and an SP1-HA prey protein in the TEAMP system (Fig. 4.7). Moreover, in vitro pulldown assays using GST-SP1flex as bait revealed a strong interaction with SP2 (Fig. 4.8). In the latter system, native, radiolabelled in vitro-translated SP2 was used, but SP1 had its transmembrane domains deleted.

While both of the aforementioned protein-protein interaction experiments have certain limitations (as discussed below), the two approaches are highly complementary. Thus, the fact that they both gave the same result, combined with the strength of the interaction seen in each case, strongly supports the idea that SP1-SP2 is a true, physiologically relevant interaction. Nonetheless, this notion could be further supported by applying other, complementary approaches such BiFC.

In the TEAMP system, although SP2-Myc is already expressed in the tissue used, it is possible that not all of the protein is fully assembled at its functional destination (see Chapter 3). If so, this could lead to the detection of non-specific interactions. For example, the intermembrane space domain of SP1-HA could interact with unassembled SP2-Myc in the IMS. Although most

chloroplast OEPs do not require import before assembly (Shi and Theg, 2013b), if SP1 were first imported (its route of assembly is currently unknown), it is possible that a non-specific interaction could occur between unassembled SP2-Myc and SP1-HA in the IMS.

In the *in vitro* system, SP1 had its transmembrane domains deleted, which were replaced with a short flexible linker, to create SP1flex. This could mean that the interaction between SP1 and SP2 occurs between SP1's cytosolic or IMS domains and SP2's N-terminal sequence or its loops between β -strands. This is feasible, and an interaction between SP1's IMS domain and its substrates was already demonstrated (Ling et al., 2012). As SP2 is mostly a β -barrel OMP, with no other major domains, and SP1 is a transmembrane protein, it is potentially feasible that their interaction occurs between their transmembrane domains. However, it would not be possible to observe such an interaction in this *in vitro* experiment, using SP1flex. Another limitation of this experiment is that, although *in vitro* translated SP2 is clearly soluble (and therefore not aggregated), and does not interact with the beads or GST non-specifically, it is out of the context of the lipid bilayer and so it may not interact with proteins in the same way that it would in the biological context. This could lead to non-specific interactions between the exposed domains of SP2 and GST-SP1flex.

4.3.2 The relationship of SP2 to the ChlorAD hypothesis

Overall, the presented evidence suggests that SP2 forms part of the hypothesised chloroplast-associated degradation pathway, and thus acts in concert with, and downstream of, SP1. In one pathway of ERAD, Hrd1 is the primary E3 ligase that recognises and ubiquitinates substrates, which it then retrotranslocates through a pore formed from its six transmembrane spans. In

this process, the ubiquitin-interacting protein and ATPase, CDC48, binds to the substrate and provides the motive force for its retrotranslocation to the cytosol for proteasomal degradation (Stolz et al., 2010; Ruggiano et al., 2014; Baldrige and Rapoport, 2016). A similar system, dubbed MAD, was found in mitochondria (Wu et al., 2016).

Membrane proteins require a mechanism of removal in addition to simply being modified with ubiquitin. This is because they evolved to sit within a lipid bilayer, and instability or misfolding cannot be counted upon to cause the protein to leave the bilayer, especially as proteins may often require degradation before the end of their useful life (Avcı and Lemberg, 2015). Given the discovery of SP1, an energy-dependent pathway must be required for the extraction of the TOC proteins. By analogy to ERAD, the steps required are: substrate identification, ubiquitination, entry into the retrotranslocon, binding of the ATPase motor, and energy-dependent retrotranslocation. SP1 fulfils the substrate identification and ubiquitination functions, and CDC48 fulfils the role of ATPase motor and energy-dependent retrotranslocation (unpublished). It is possible that SP1 could provide the missing piece, fulfilling the role of retrotranslocon as well. However, SP1 is not structured like Hrd1, which has six transmembrane spans, and so SP1 would likely need to form a homomultimer in order to do so. The discovery of, and information gathered on, SP2 suggests an alternative possibility, and identifies SP2 as an ideal candidate to fulfil the role retrotranslocon in this system.

It was found that SP2-Myc specifically pulls down Toc33, Toc75-III and Toc159 in co-IP studies in Chapter 3; unlike the case of SP1-HA, these are the native proteins from the chloroplast. In contrast, Tic110 and Tic40 were not pulled down by SP2-Myc, neither were the TOC proteins pulled down by the

beads themselves. This interaction therefore indicates that SP2 can associate with the TOC proteins, and is consistent with its proposed role as the retrotranslocon for ChlorAD.

Moreover, SP2-Myc was able to pulldown HA-CDC48 in the TEAMP system (Fig. 4.7), although the enrichment of HA-CDC48 was not as strong as that of SP1-HA. Assuming that this is a true interaction, the ChlorAD hypothesis predicts that it would be mediated by SP2's binding of ubiquitinated substrates, to which CDC48 would also be bound via its ubiquitin-binding domain and/or adaptor proteins (Wolf and Stolz, 2012).

In summary, SP2 is involved in the negative regulation of TOC proteins, and, due to its interaction with SP1, both genetically and physically, SP2 is implicated in the UPS-dependent degradation of TOC proteins via the ChlorAD mechanism. The presented data are consistent with the notion that SP2 acts as the retrotranslocon in the hypothesised ChlorAD pathway.

4.4 Figures

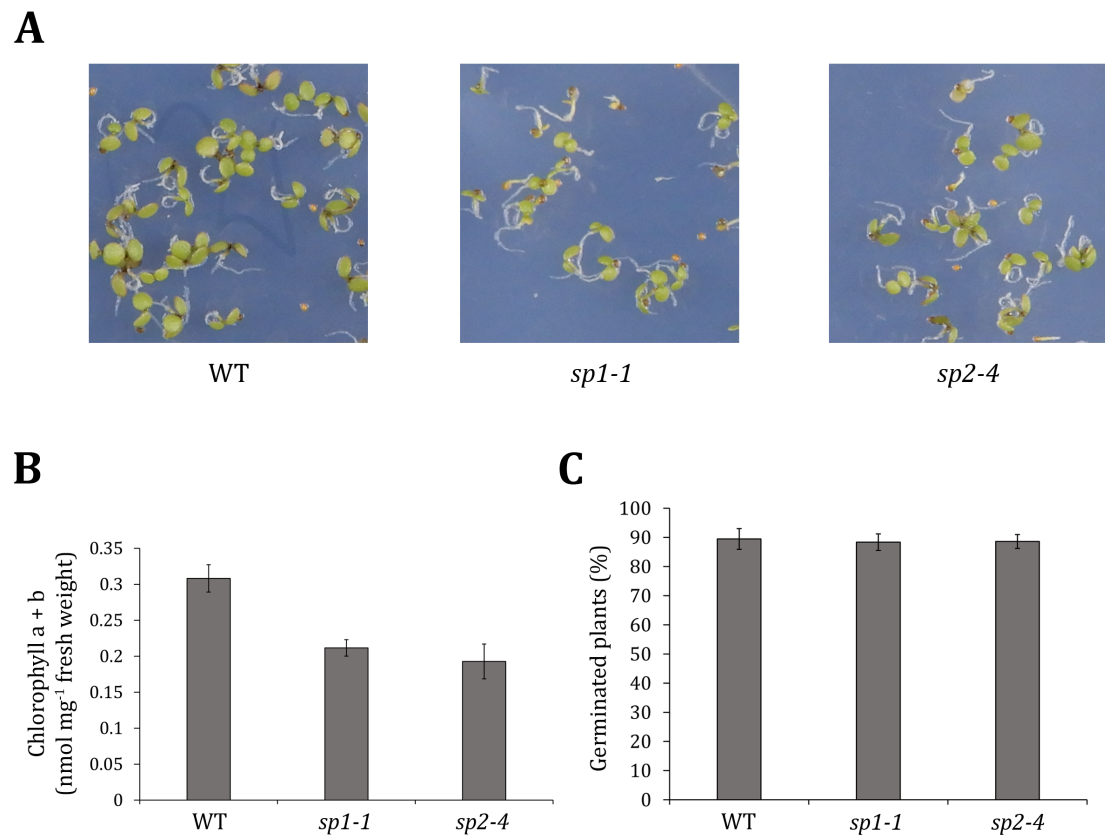


Figure 4.1. Analysis of salt stress on *sp1* and *sp2* plants. Seeds of wild type, *sp1* and *sp2* were germinated on MS media containing 150 mM NaCl, and grown for 14 days. (A) Representative images of plants, and (B) quantification of chlorophyll extracted with DMF from whole, 14-day-old seedlings. (C) Assessment of the number of seeds germinated. Error bars represent SD (n = 6).

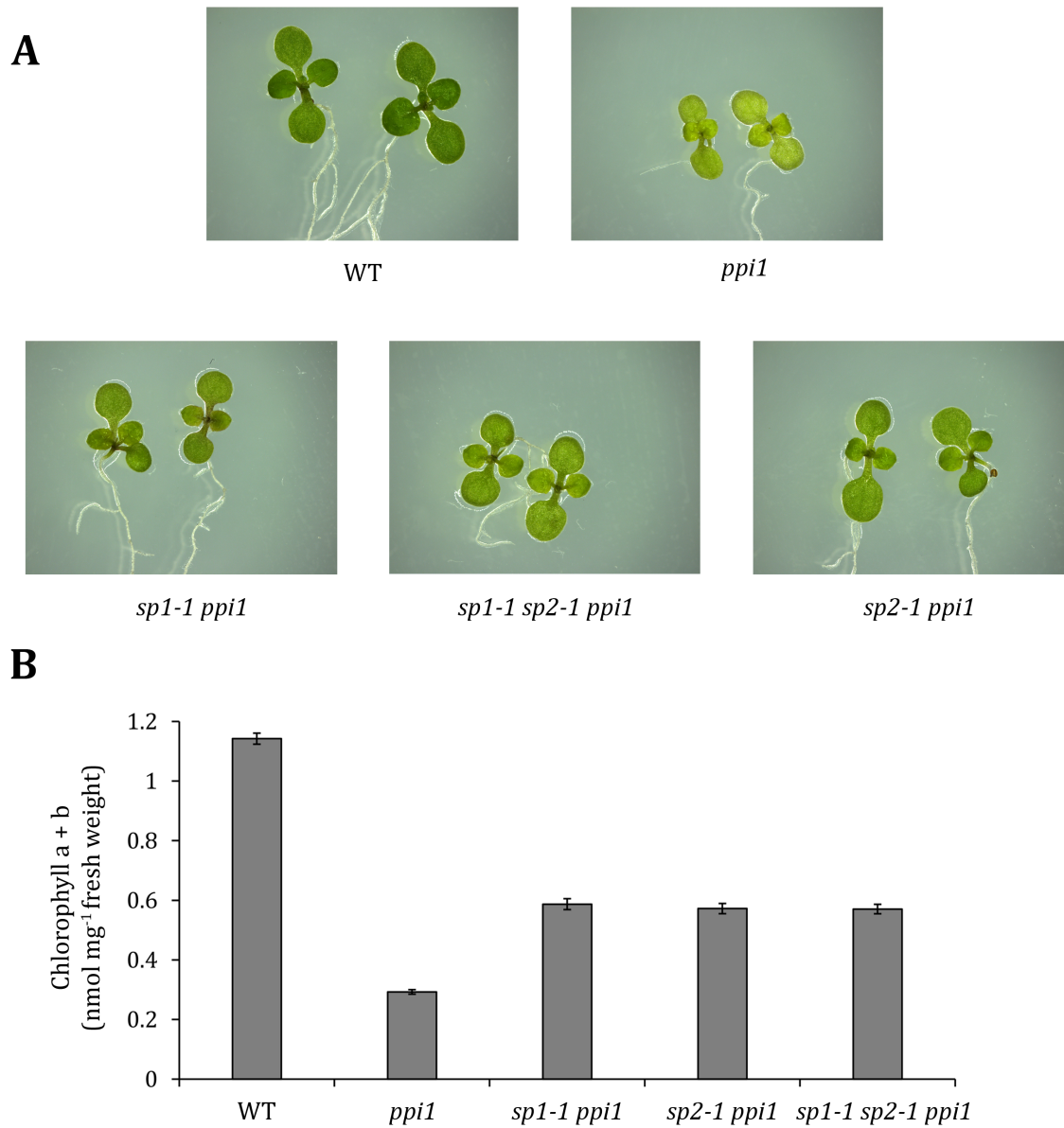


Figure 4.2. Phenotypic analysis of *sp1 sp2 ppi1* plants. (A) Representative images of, and (B) quantification of chlorophyll extracted with DMF from, whole 10-day-old *sp1 sp2 ppi1* plants and appropriate controls. Error bars represent SEM (n = 4).

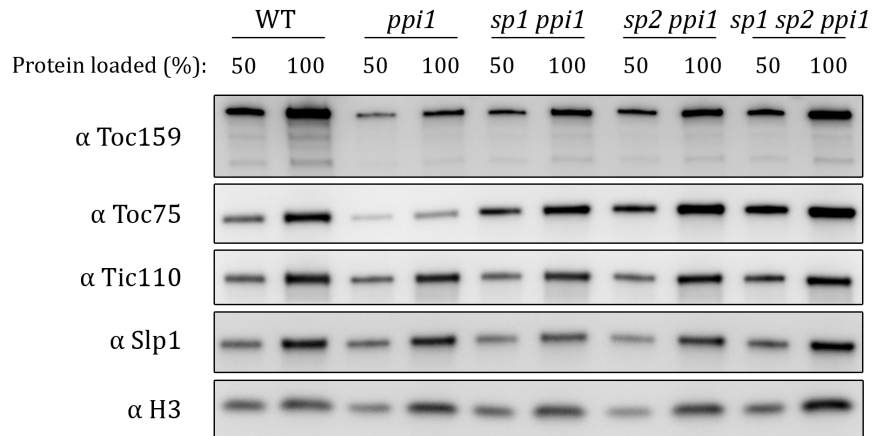
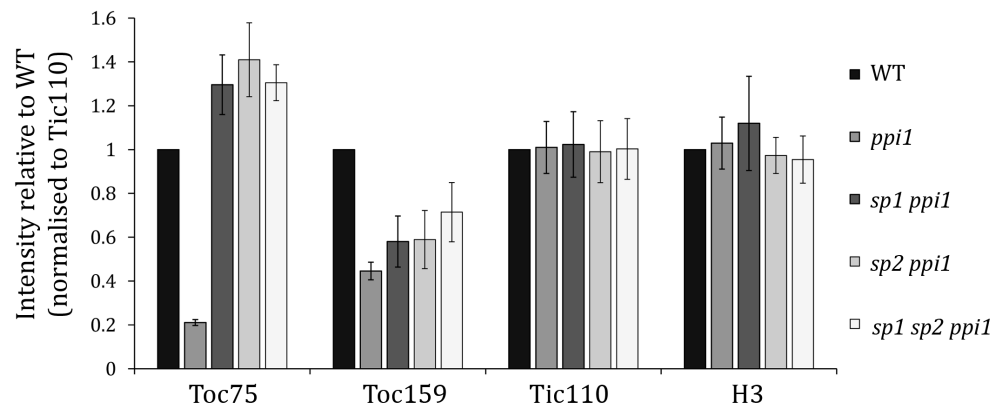
A**B**

Figure 4.3. Semi-quantitative western blot analysis of chloroplast protein abundance in *sp1 sp2 ppi1* mutant plants. (A) Equal amounts of total protein extracts from whole, 10-day-old seedlings were analysed using the indicated antibodies, and representative western blot images are shown. (B) Semi-quantitative analysis of western blot images corresponding to four or five biological replicates. The data for each protein were normalised to corresponding data for mitochondrial Slp1, and are presented as a ratio to the wild type value for each respective protein. Error bars represent SEM. $n = 4$ to 5 .

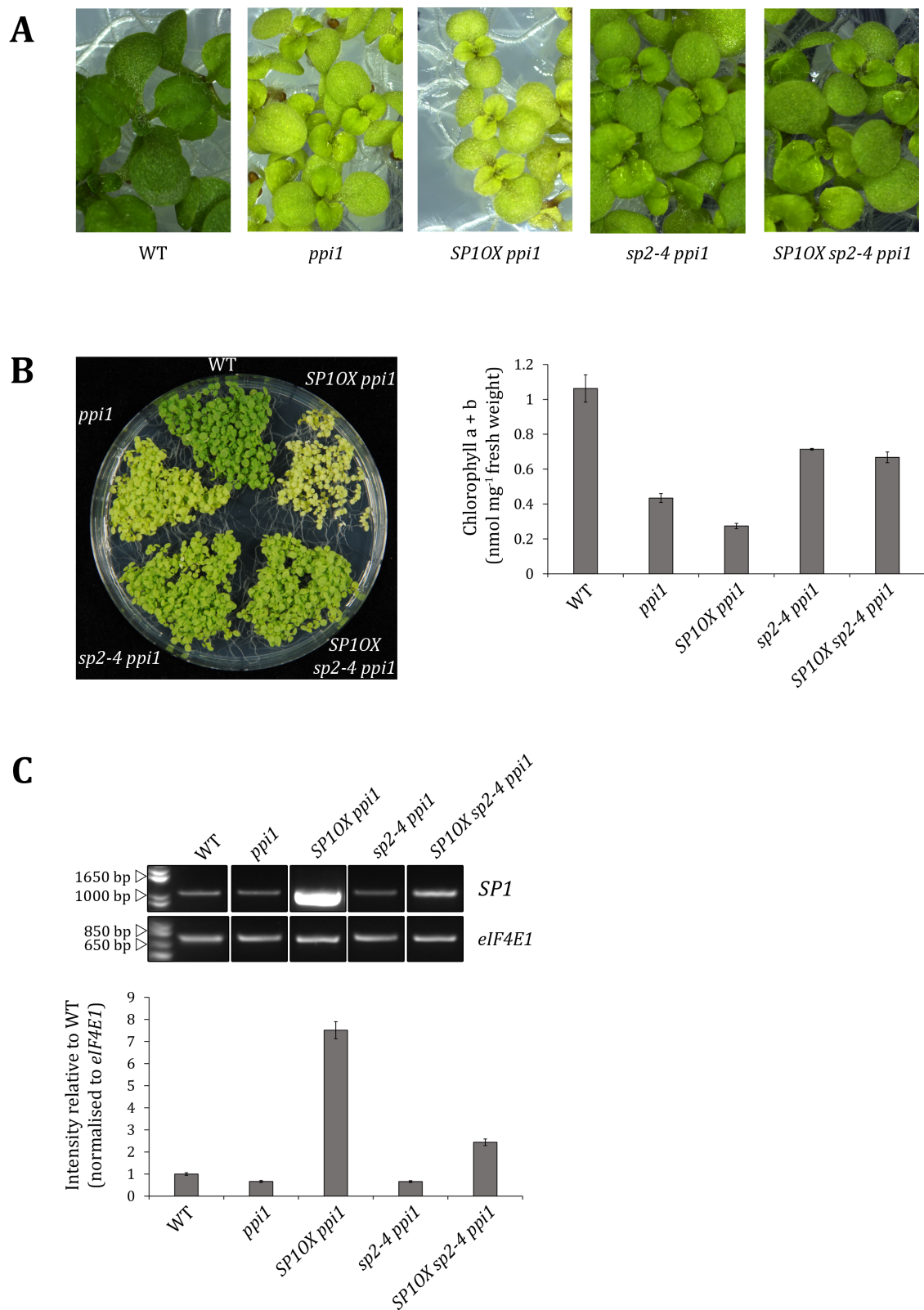


Figure 4.4. Analysis of the ability of *sp2* to suppress *SP10X ppi1*. (A) Representative images of 10-day-old plants. (B) Image of plate of 10-day-old plants together on a single plate and quantification of chlorophyll extracted with DMF from whole 10-day-old seedlings. (C) Semi-quantitative RT-PCR analysis of

SPI gene expression, using *eIF4E1* as a control. A representative gel image (top), alongside quantification of four biological replicates (bottom). Error bars represent SEM. n = 4.

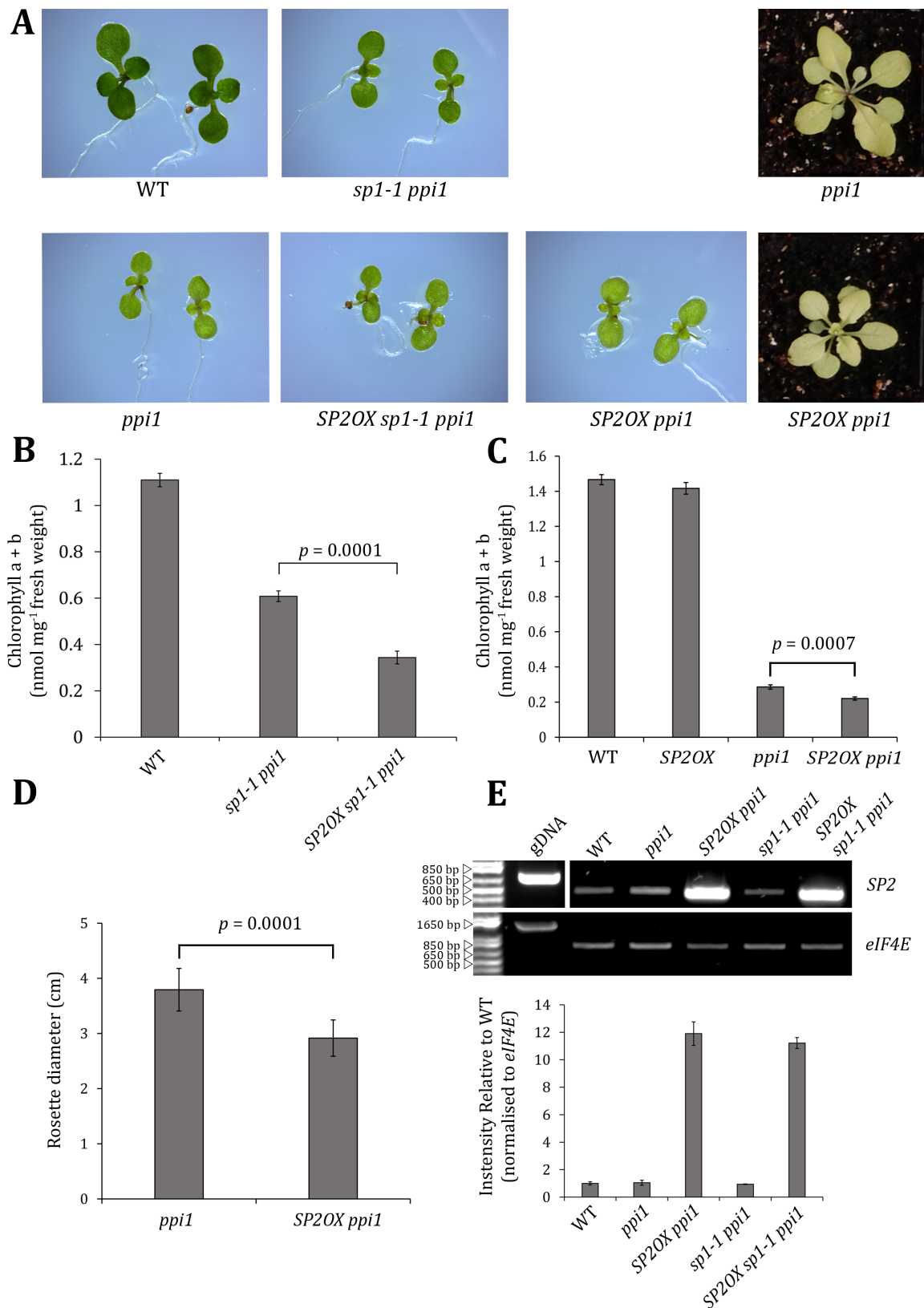


Figure 4.5. Analysis of the ability of *sp1* to suppress *SP2OX ppi1*. (A) Representative images of 10-day-old plants (left), and 28-day-old-plants (right). (B and C) Quantification of chlorophyll extracted with DMF from whole 10-day-old *SP2OX sp1-1 ppi1* plants (B), and chlorophyll quantified using SPAD-502

meter measurements converted to chlorophyll concentration (Ling et al., 2011) from 28-day-old *SP2OX ppil* plants (C). Error bars represent SEM (n = 4). (D) Rosette diameter measurements from *SP2OX ppil* and *ppil* plants, as measured at the widest point. Error bars represent SD (n = 9 and 10 respectively). (E) Semi-quantitative RT-PCR analysis of *SP2* gene expression, using *eIF4E1* as a control. A representative gel image (top), alongside quantification of four biological replicates (bottom). Error bars represent SEM. n = 4. *p* values were derived from a t-test of the means of the indicated genotypes.

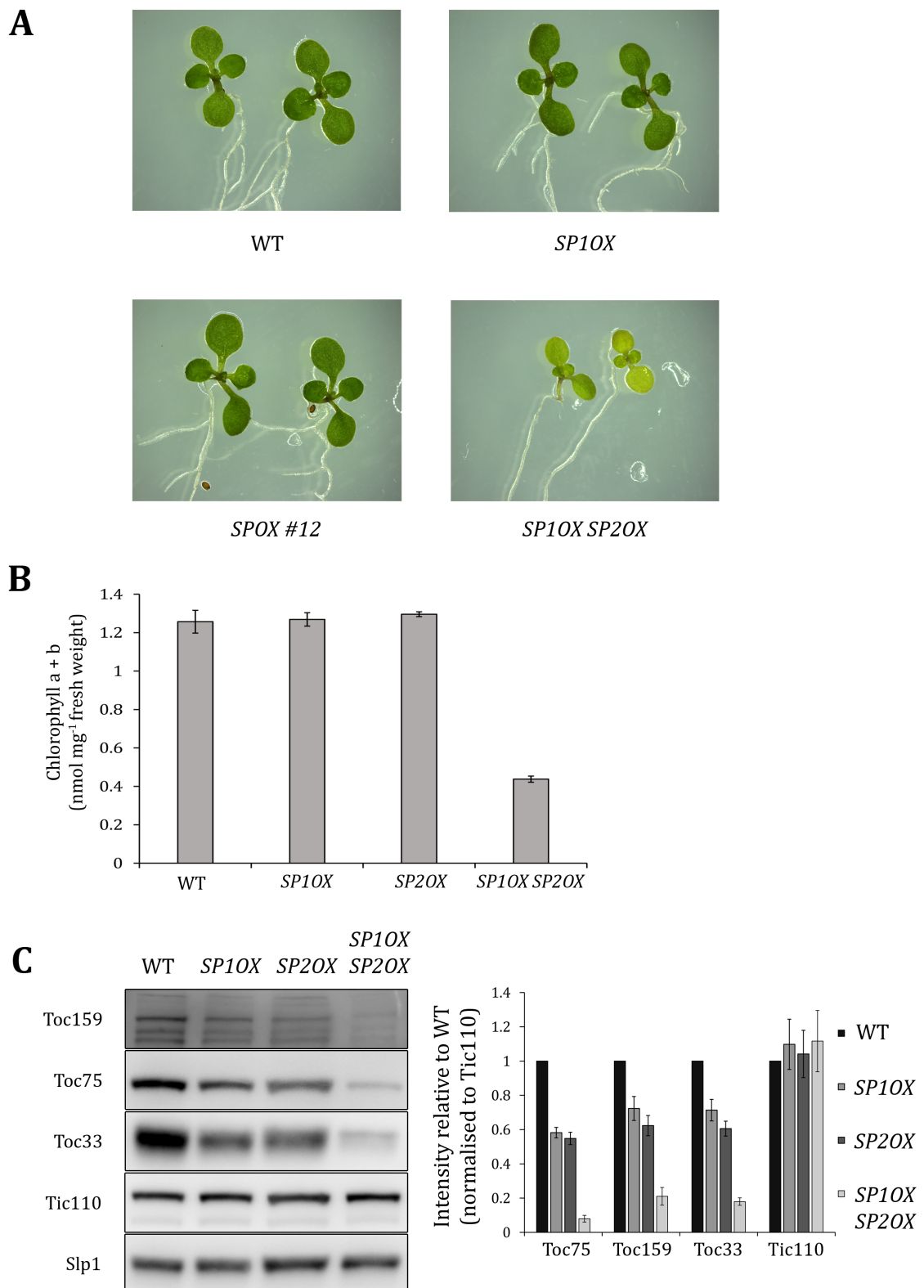


Figure 4.6. Phenotypic analysis of *SP10X SP20X* plants. (A) Representative images, and (B) quantification of chlorophyll extracted with DMF from whole 10-day-old *SP10X SP20X* plants with relevant controls. (C) Equal amounts of total protein extracts from whole, 10-day-old seedlings were analysed using the

indicated antibodies, and representative western blot images are shown (Left). Semi-quantitative analysis of western blot images corresponding to four biological replicates (Right). The data for each protein were normalised to corresponding data for Slp1, and are presented as a ratio to the wild type value for each respective protein. Error bars represent SEM. n = 4.

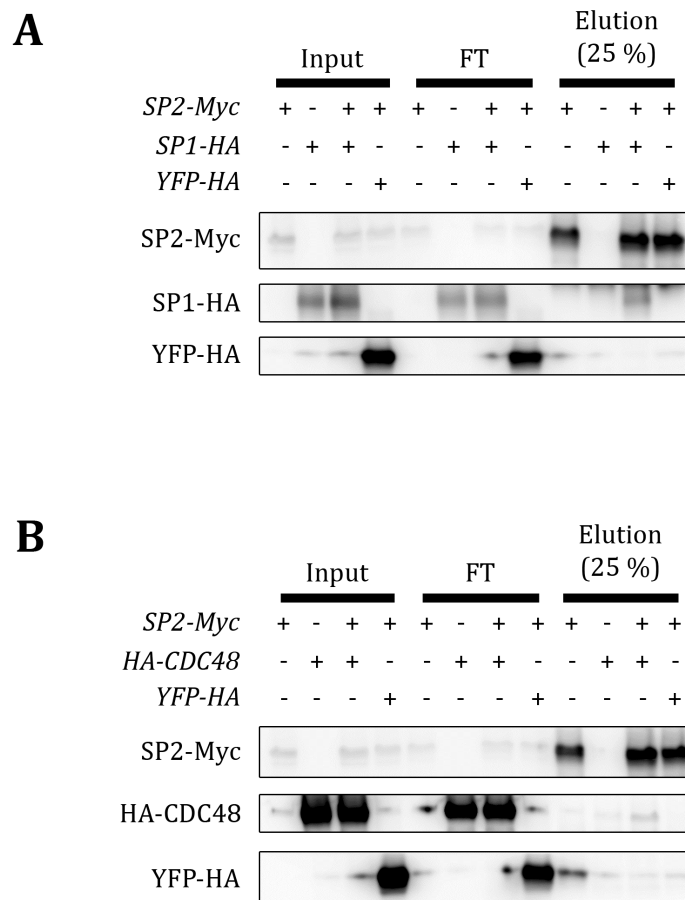


Figure 4.7. Co-immunoprecipitation of SP1-HA and HA-CDC48 from *Sp2-Myc sp2-4* mesophyll protoplasts. (A) SP1-HA or YFP-HA were transiently expressed in mesophyll protoplast from *SP2-Myc sp2-4*, or wild type plants. Consequently, co-IP was performed using SP2-Myc as bait. (B) HA-CDC48 or YFP-HA were transiently expressed in mesophyll protoplast from *SP2-Myc sp2-4*, or wild type plants. Consequently, co-IP was performed using SP2-Myc as bait. FT = Flow through; +/- indicates which constructs were expressed in each sample.

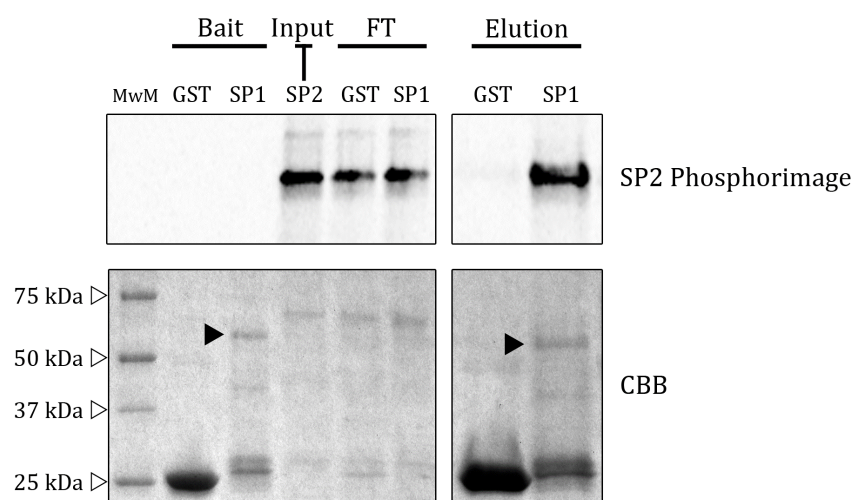


Figure 4.8. In vitro pulldown of SP2 by SP1. Pulldown of in vitro translated radiolabelled SP2 was performed, using GST-SP1flex or GST, each bound to glutathione agarose beads, as bait. The prey was incubated with the bait, the beads were washed, and then eluted with free reduced glutathione. The experiment was analysed by SDS-PAGE and phosphorimaging. Black arrows indicate GST-SP1flex at 66 kDa. FT = Flow through; CBB = Coomassie brilliant blue.

Chapter 5

General discussion

5. General discussion

5.1 The context of this work

Plastids are diverse and complex organelles. Therefore, their biogenesis, functions and homeostasis, and the regulation of these processes, are also complex (Jarvis and López-Juez, 2013). A vital aspect of plastid biogenesis and functioning is protein import into chloroplasts. Over the past two decades, much progress has been made in defining the machinery involved in this elaborate molecular process (Shi and Theg, 2013b; Oh et al., 2015). The overall view of the process is well outlined. We know that most plastid preproteins are nucleus-encoded and synthesised in the cytosol, and that there is a requirement for these proteins to be targeted and chaperoned to the chloroplast (Lee et al., 2014; Richardson et al., 2014). These proteins are then transported across the envelope membranes in an energy-dependent process, which may involve stromal chaperones and which ends in the cleavage of each preprotein's transit peptide by stromal processing peptidase (SPP), before they assume their functional role or are transported further (Shi and Theg, 2013b).

Further mechanistic details of the import process are still lacking. For example, with regard to the initial process of plastid preprotein targeting, while many properties of transit peptides have been described, precise knowledge on which of these properties allow specific chloroplast targeting remain elusive (Li and Teng, 2013). A number of associated cytosolic chaperones as well as potential interactors at the chloroplast OEM, which aid in the initial targeting, have also been discovered, but this list may not be exhaustive and the exact roles of these components, and their mechanistic action in relation to targeting, are not fully

realised (Flores-Pérez and Jarvis, 2013). At the chloroplast envelopes, there are TOC and TIC complexes, which mediate preprotein transport across the membranes, and into the chloroplast (Shi and Theg, 2013b). The core components of the TOC have been defined, and we have some understanding these TOC components and their functional relevance, but we only know the properties of some of these components in depth. Moreover, the exact molecular functions of some of the best characterised components are largely unknown; for example, the functional relevance of the GTPase domains of Toc33 and Toc159 is still debated (Shi and Theg, 2013b). There are also other, less characterised TOC components, such as Toc64, and newly-proposed components such as, most recently, a kinase called Koc1 (Sommer et al., 2013; Zufferey et al., 2017). This indicates that knowledge concerning the TOC complex is still immature. The TIC machinery is arguably even less well characterised. Although there are many components which are known to be critical for transport across the inner-envelope of chloroplasts, debate is ongoing regarding the number of potential TIC complexes and which TIC components may be a part of which complex, and what the exact function of each component is (Nakai, 2015).

This lack of in-depth knowledge on the TOC complex, and more generally on the molecular mechanisms underpinning chloroplast biogenesis, are what spurred the development of a mutagenesis screen which yielded *sp1* and *sp2*. The discovery of these two proteins has opened up a different angle on the TOC complex, its regulation, and the plant's ability to modulate protein import into plastids.

The *suppressor of ppil locus 1* was characterised by Ling et al. (2012), and this work revealed that SP1 is part of a molecular mechanism which could label

components of the TOC complex for degradation by the cytosolic 26S proteasome. This was an interesting observation, as previously the UPS had not been shown to specifically target mature proteins of the chloroplast. Furthermore, the turnover of the TOC components had not been specifically studied previously. This study provided a new opportunity to shift from studying the mechanisms of protein import, the functions of the protein import components, and the interactions of plastid preproteins, towards studies on how these components are manipulated by the wider regulatory network of the cell in order to effectively regulate and manipulate plastids during their biogenesis and interconversions, and during stress conditions.

5.2 Summary of key findings

In this project, SP2 was identified and characterised. SP2 appears to also operate in the realm of TOC protein degradation. On its own, this in itself would provide an expansion on our knowledge concerning the regulation of the TOC components. However, there is additionally evidence for SP2's interaction with SP1, which leads to the interesting prospect that SP1 and SP2 act in a single pathway.

It was found that SP2's role in chloroplast biogenesis was directly related to protein import. By genetic interaction studies, it was found that SP2 is related specifically to protein import across the OEM, and in particular the TOC proteins, as mutants of the three main TOC proteins which could be suppressed by *sp2*. Molecular evidence supported this interpretation, as the mutation of *sp2* caused the accumulation of these TOC proteins, while the overexpression of *SP2* led to the depletion of TOC proteins. Moreover, this regulation must occur at the post-transcriptional level, as there were no effects on TOC transcripts. Furthermore, a direct interaction was observed between SP2 and the three main TOC proteins, via co-IP using SP2-Myc as bait.

At this point, it was abundantly clear that these observations were the same as those for the previously-characterised and published SP1 protein. Thus, the potential interaction between SP1 and SP2 was studied. Firstly, more specific functional similarity was tested. As *sp1* was previously reported to exhibit sensitivity to high salinity, this was tested for *sp2* also, and it was found that *sp2* was equally sensitive as *sp1* to the high-salt media. Moreover, there was no phenotype additivity when both *sp1* and *sp2* were mutated in *ppi1*, and this was

reflected at the protein level. In contrast, the overexpression of both SP1 and SP2 caused a synergistically severe phenotype, reminiscent of *ppi1*, reflecting an enhanced depletion of TOC proteins in these plants.

The recent findings of others in the group that SP1 could interact with CDC48 to effect TOC protein degradation had led to an interesting hypothesis – namely, that there is a specialised degradation mechanism operating at the chloroplast envelope (unpublished). This would be analogous to the already well-known and characterised ERAD and MAD pathways, and could potentially employ similar (or homologous) components (such as CDC48) and mechanisms.

5.3 Critical appraisal of the role of SP2 in ChlorAD

5.3.1 Assessment of the ability of SP2 to participate in ChlorAD

ERAD and MAD solve an important molecular problem associated with organelle membranes. The ERAD and MAD substrates, like the TOC proteins of chloroplasts, can be observed to be degraded by the UPS (Taylor and Rutter, 2011; Wolf and Stolz, 2012; Bragoszewski et al., 2013). However, there is a significant problem posed by the organellar membrane, which is a barrier to facilitating this process, as UPS components do not exist inside these organelles (and neither do they exist inside the chloroplast). In ERAD there are luminal substrates, and in MAD there are IMS substrates, and so of course these need to be transported to the cytosol. In addition, there exist integral membrane proteins in the ER and mitochondria which are UPS substrates (like the TOC proteins in the chloroplast), and these too must require removal from the membrane prior to their degradation (Taylor and Rutter, 2011; Avci and Lemberg, 2015). It has been found that the ERAD and MAD systems solve both of these problems by utilising a membrane channel called a retrotranslocon, to allow substrate transport from the non-cytosolic side of the membrane, and to assist in the removal of membrane protein substrates (Bragoszewski et al., 2015; Baldridge and Rapoport, 2016). The motive force for such translocation comes from the AAA+ ATPase, CDC48, in the cytosol, which interacts with ubiquitin conjugated to the substrate protein via interactors such as Ubx proteins in ERAD and Doa1 in MAD (Wolf and Stolz, 2012; Wu et al., 2016).

Given that the TOC proteins are integral to the OEM, and the chloroplast contains no UPS, a similar molecular problem exists in chloroplasts, and so a

ChlorAD mechanism was hypothesised. CDC48 has now been associated with the process (unpublished), but the other critical role of retrotranslocon remained to be elucidated. SP2 presents itself as a likely candidate for this role (Fig. 5.1).

Given the discovery that SP2 is a β -barrel OMP, it would seem possible for SP2 to fulfil a role in OMP extraction. Functions of related Omp85 proteins include membrane protein biogenesis, and transport across the outer-membranes of gram negative bacteria, mitochondria, and chloroplasts (Misra, 2012). Data presented in this thesis, along with those in previous reports, indicate that SP2 forms a 16-stranded β -barrel homologous to OEP80, which has channel properties and is localised to the chloroplast (Hsueh et al., 2017). Thus, structurally, SP2 would appear to have the ability to transport proteins, and it is conceivable that an Omp85 protein could act to facilitate removal of proteins, rather than assembly or transport.

But how, in the physical environment of the membrane, could SP2 act to remove and transport its substrate proteins? Retrotranslocation of a membrane-bound protein has been reported, but this was by a multi-transmembrane span protein, Hrd1, rather than a β -barrel protein (Baldrige and Rapoport, 2016). Structural reports of Omp85 proteins which function in the assembly of proteins into the bacterial outer membrane show that the first and last β -strands of the barrel have open and closed conformations, which was observed in molecular dynamics simulations (Noinaj et al., 2013). Such lateral opening of the channel would permit an entry pathway for substrate β -strands or transmembrane domains (Noinaj et al., 2013). This idea was tested by engineering disulphide bridges between the first and last β -strands of the barrel, and this ablated Bama function, indicating that lateral opening is essential for function (Noinaj et al.,

2014). Lipid bilayer disruption and membrane thinning were also observed, and suggested as possible mechanisms which might assist membrane protein assembly (Noinaj et al., 2013). Potential consequences of this membrane-thinning are observed when BamA is assembled into proteoliposomes at high concentrations (25:1, lipid to protein ratio), and causes curious structural defects to the proteoliposome, visible by SEM, including notches, and in extreme examples, large, angular invaginations (Sinnige et al., 2014).

Such lipid bilayer disruption was interpreted to be important in enabling structures of the substrate protein to more easily pass from the β -barrel lumen into the membrane environment as it is assembled (Noinaj et al., 2013; Noinaj et al., 2014). It is conceivable that these properties could be taken advantage of for protein extraction, for instance, so that a substrate protein associated with SP1 could be passed into the β -barrel lumen to facilitate transport. Moreover, one can speculate that the local disruption of the lipid environment by SP2 could itself be sufficient to allow for protein extraction, provided that the substrate can be brought into a sufficiently close association with SP2.

Further analysis of the SP1 and SP2 interaction was required before such a speculative interpretation could be supported. Genetic studies were conducted where it was assessed whether *sp2* could suppress *SP1OX ppil*, or whether *sp1* could suppress *SP2OX ppil* (Figs. 4.3, 4.4). The hypothesis was that if these two proteins are dependent upon each other, then knocking out the function of one of the components, would disable the phenotypic severity conferred by overexpression of the other component in *ppil*. This seemed to be the case for *SP1OX sp2 ppil*, although the degree of SP1 overexpression was low (Fig. 4.3). However, it was found that *sp1* did not suppress *SP2OX ppil*, meaning that SP2

can act in the absence of SP1 to remove TOC proteins (Fig. 4.4). This result was not as predicted, but it is not inconsistent with ChlorAD. It does, however, suggest that the operation of the system, and of the different components involved, is more complex than envisaged. Given that there are already two known chloroplast-localised SP1 homologues, it seems feasible that there are other E3 ligases (or other factors entirely) that work together with SP2, and that these enable SP2 to function in the absence of SP1; thus, *SP2OX* would behave epistatically in relation to *sp1*, with SP2 acting downstream of SP1 in the pathway.

To corroborate the interpretation of the genetic interaction data, there is considerable support for the interaction between SP1 and SP2 was provided by the physical protein interaction data. SP2-Myc was able to pulldown SP1-HA in co-IP experiments (Fig. 4.6). In a reciprocal scenario, GST-SP1flex was able to pulldown in vitro translated SP2 (Fig. 4.7). Both experiments showed a strong and clear association between the SP proteins, whereas the controls were negative indicating specificity.

Providing further consistence with the proposed function of SP2 with SP1, SP2-Myc was able to pulldown CDC48 (Fig. 4.6), which is an interesting result because in view of the ChlorAD model, as the association would be via SP1 or a substrate protein (Fig. 5.1). Both interactions are consistent with the ChlorAD model.

The evidence provided for the participation of SP2 in the ChlorAD pathway is compelling, though more experimentation is still needed to provide a more robust interpretation of the role of SP2, as well as to gain further mechanistic understanding of ChlorAD. These would include, direct tests of ability of SP2 to facilitate protein removal from the chloroplast OEM, experimentation to test

whether SP2 influences UPS degradation of TOC proteins, as well as a experimentation to test potential interactions of SP2 with SPL proteins.

5.3.2 Alternative interpretations for SP2 function

There are other possible interpretations notwithstanding. One option is that SP2 could be a part of a specialised assembly pathway for SP1 (or perhaps other negative regulatory machinery in the OEM). However, the ability for SP2 overexpression to cause reversion of *sp1-1 ppil* to the *ppil* phenotype would not be consistent with such a function as an assembly factor for SP1 (but may still for other E3 ligases or unknown factors). Testing localisation of transiently-expressed fluorescently-tagged SP1 or SP1-like proteins in *sp2* plant cells would be a step towards resolving this.

Another possible interpretation is that SP2 has a separate role that is related to the UPS degradation of TOC components, but which can act separately to that process. For example, SP2 could associate with a TOC complex in order to disrupt and disperse the complex (perhaps caused by a lack of POTRA domains to anchor certain components), which would then mean there is a greater potential for these TOC components to rearrange and also interact with the E3 ligases. This interpretation would be consistent with the data reported here, including the apparent inability of *sp1* to suppress *SP2OX ppil*, and the observation that *SP2OX* does not severely accentuate the *ppil* phenotype (Fig. 3.4). With regard to the latter result, the lack of Toc33 (in *ppil*) could mean that there are already no stable TOC complexes and so the E3s already have sufficient access to substrates, whereas in *ppi2-3*, where SP2 overexpression does accentuate the phenotype (Fig. 3.4), there could still be relatively intact TOC complexes, because Toc159 is still present, though truncated. This hypothesis may be tested by assessing TOC

complex topology in SP2 overexpressors and mutants, for example. In this scenario, the identity of the retrotranslocon required for the SP1 pathway would remain in question.

Alternatively, SP2 may be part of an SP1-independent mechanism to regulate TOC proteins, and in this case, there may be few clues as to SP2 function in the data presented in this thesis, though retrotranslocon could still be an option (however, this would mean that the identity of the retrotranslocon working with SP1 would remain unknown). This scenario would raise the question – Why would there exist two independent, unrelated mechanisms to regulate the same components at the chloroplast OEM?

5.4 Further perspectives

5.4.1 The influence of ChlorAD on specific plastid protein import pathways

When SP1 was first described, its overexpression was reported to accelerate the rate of two developmental stages, the greening of etiolated seedlings, and leaf-senescence (Ling et al., 2012). The greening of etiolated seedlings is critically centred on the etioplast to chloroplast transition, while senescence involves the chloroplast transition to gerontoplast, and partial degradation of the plastid (Jarvis and López-Juez, 2013; Xie et al., 2015). This was interpreted as SP1 selectively degrading the dominant TOC complex-type of those plastids, the complex containing Toc132/120 in etioplasts, or the complex containing Toc159 in chloroplasts, which would cause the import of preproteins favoured by that complex could be halted, while allowing space in the OEM for an alternative TOC complex-type to assemble in its absence (Kessler, 2012; Ling et al., 2012; Jarvis and López-Juez, 2013). The more efficient assembly of the new TOC complex-type would expedite the import of that complex's cognate preproteins required to cause the plastid interconversion. This might also enable other components to assemble, or alter the OEM interactions with cytosolic components on other ways.

Consistent with this predicted targeting of specific TOC components by SP1, there was evidence that *SP1OX*, and *sp1-1*, affected the abundance of Toc34, Toc132 and Toc120 during de-etiolation, with *sp1-1* plants having a markedly lower Toc159 to Toc132 ratio than wild type, and *SP1OX* plants having a greater Toc159 to Toc132 ratio than wild type (Ling et al., 2012).

Given the ChlorAD model, it would be expected that ChlorAD, including SP2 would facilitate the removal of these components during this process and so we would expect that mutants or overexpressors of SP2 could cause similar phenotypes, and effects to TOC proteins, as those recorded for SP1. So far only the abundance of Toc34 in *sp2* and *SP2OX* plants has been recorded, and this was in 28-day-old leaf tissue, where differences in Toc34 abundance were not observed, however this may not be the most appropriate test. Therefore, the contribution of SP2 to degradation of TOC complexes during plastid interconversions remains to be assessed.

It is also worth noting that there are many different TIC proteins, some of which have multiple isoforms like the TOC proteins, which could form different TIC complexes (Jarvis and López-Juez, 2013; Kikuchi et al., 2013; Shi and Theg, 2013b; Nakai, 2015). Precisely whether different TIC complexes are involved in the import of particular preproteins, the stage of protein import at which different TIC components are involved, and how such TIC complexes may interact with the TOC complexes is not fully understood (Jarvis and López-Juez, 2013; Shi and Theg, 2013b; Nakai, 2015). But given the variety of TIC components, like the TOC proteins, the composition and abundance of different TIC complexes could be important for protein import. As ERAD and MAD systems extend beyond the membranes in which they are seated and into the organelle beyond (Ruggiano et al., 2014; Bragoszewski et al., 2015), one could speculate that ChlorAD might do the same, and that the TIC components could be such targets. However the observed abundance of Tic40 and Tic110 in the SP mutants and overexpressors remains unchanged and so would render this an unlikely scenario (Fig. 3.8; Ling et al., 2012; Ling and Jarvis, 2015b), however as

Tic20 is also argued to be a TIC protein conducting channel, both instead of or in addition to Tic110, and has multiple isoforms as is the case for the TOC proteins, it may be a better candidate for investigation (Kikuchi et al., 2013; Shi and Theg, 2013b; Nakai, 2015).

5.4.2 Questions on the evolutionary origins of ChlorAD

Assuming that the theory that both SP1 and SP2 are working as part of the same pathway is true, an interesting evolutionary story presents itself. SP1 and its homologue, SPL2, are present throughout all land plants (Fig. 5.2), and these homologues are likely true orthologues, as when using Arabidopsis SP1 as a query, the BLAST results for these species reciprocally return SP1 from Arabidopsis. However, SP2 is only present in angiosperms (Figs. 2.8, 2.9). This raises a question as to the function of SP1 in the earlier divergent plant lineages. If the function of SP1 is the same in those species, then what would take the place of SP2 in the ChlorAD mechanism? Alternatively, did SP1 neofunctionalize such that the ChlorAD mechanism is specific to angiosperms?

In response to the first question, it may be suggested that there are ample candidates to fill the role. In the plastid membrane there are other transporters most notably, close relatives of SP2, OEP80, and Toc75-III (Töpel et al., 2012). Such proteins could fulfil transport, assembly and removal, as is the case for Tom40 in MAD (Bragoszewski et al., 2015). Partial degradation of the organelle (Ishida et al., 2014; Xie et al., 2015), or even more specifically, degradation of parts of the membrane (Breuers et al., 2012) could also provide a similar solution. Evolution of SP2 therefore could have facilitated a more efficient or more versatile pathway in the angiosperms.

Part of the answer to the second question could lie in the structure, function and dynamics of chloroplasts in these early divergent lineages of plants (and perhaps algae too). Evidence suggests that moss plastids are also morphologically dynamic, contain chromoplasts and amyloplasts, and could function in gravitropism (Schwuchow et al., 2002; McCuaig et al., 2015), and analysis of TOC homologues encoded in the *P. patens* genome would suggest that mosses could also have a diversity of TOC configurations (Kalanon and McFadden, 2008). Despite this, these organisms lack many of the tissues that angiosperms have, and therefore they would also lack the same plastid interconversions. Therefore, there may be little pressure to drive the evolution of a versatile system that could change TOC components, and so SP1 may fulfil a different role in these plant lineages. However, even if these plants were to have a lack of necessity for plastids to interconvert, these plants must still have to deal with stress conditions, and so this raises the question of whether mechanisms which manipulate protein import into plastids play a role in stress responses in these species, or whether this adaptation in angiosperms (Ling and Jarvis, 2015b) is a unique innovation of the theorised ChlorAD mechanism. A definitive answer to both of these questions therefore remains to be resolved.

5.5 Figures

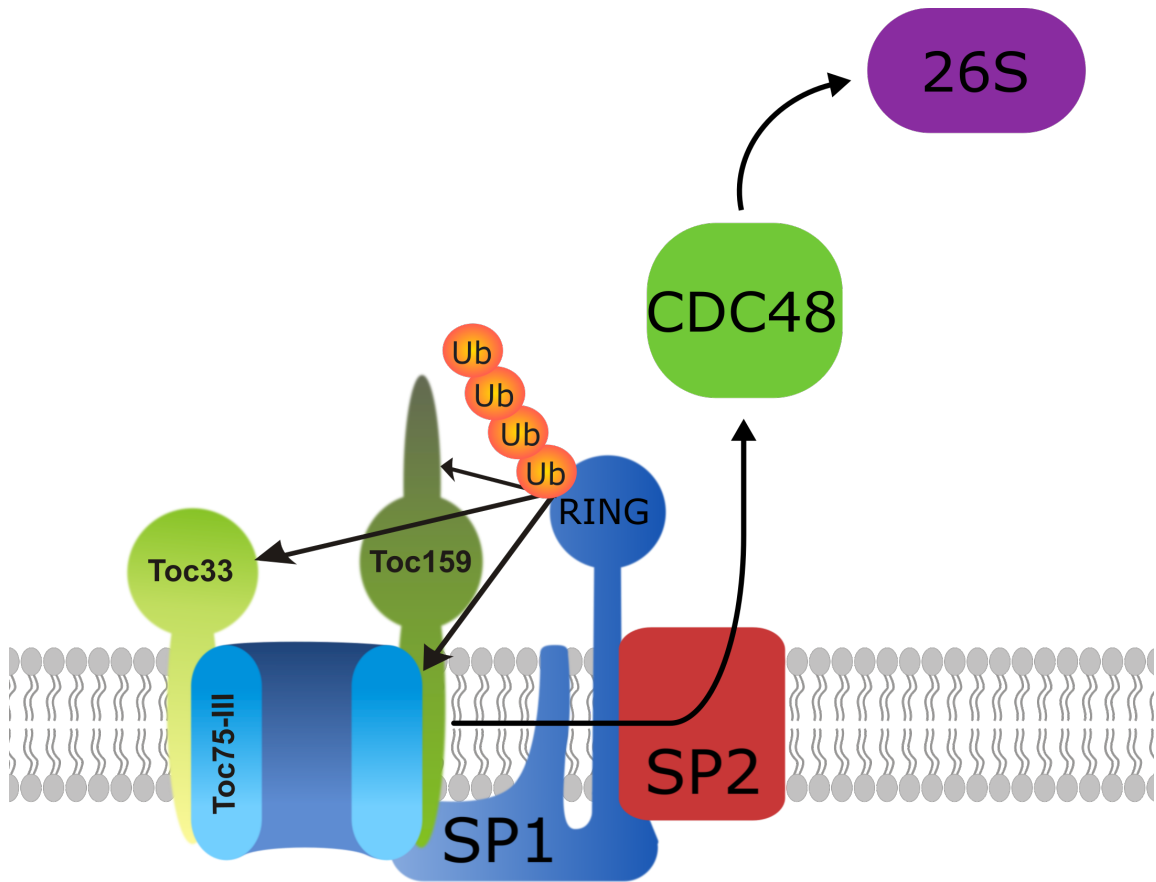


Figure 5.1. Model of the hypothesised chloroplast-associated degradation (ChlorAD) pathway at the outer-envelope membrane of chloroplasts. SP1, the E3 ligase, associates with its substrate TOC proteins and SP2, via its IMS domain. The RING domain of SP1 facilitates transfer of polyubiquitin to the substrate proteins. The substrates associate with the retrotranslocon, SP2, which facilitates their extraction from the membrane by the ATPase, CDC48. Once in the cytosol, substrate proteins are degraded by the 26S proteasome. Arabidopsis TOC proteins are shown, and are labelled by convention according to their molecular weight in kilodaltons; SP1 = Suppressor of *ppi1* locus 1; SP2 = Suppressor of *ppi1* locus 2; RING = Really interesting new gene, an E2 interacting domain; CDC48, Cell division cycle protein 48, a cytosolic AAA+ ATPase, and ubiquitin interacting protein; 26SP = the cytosolic 26S proteasome, a ubiquitin-gated proteolytic complex.

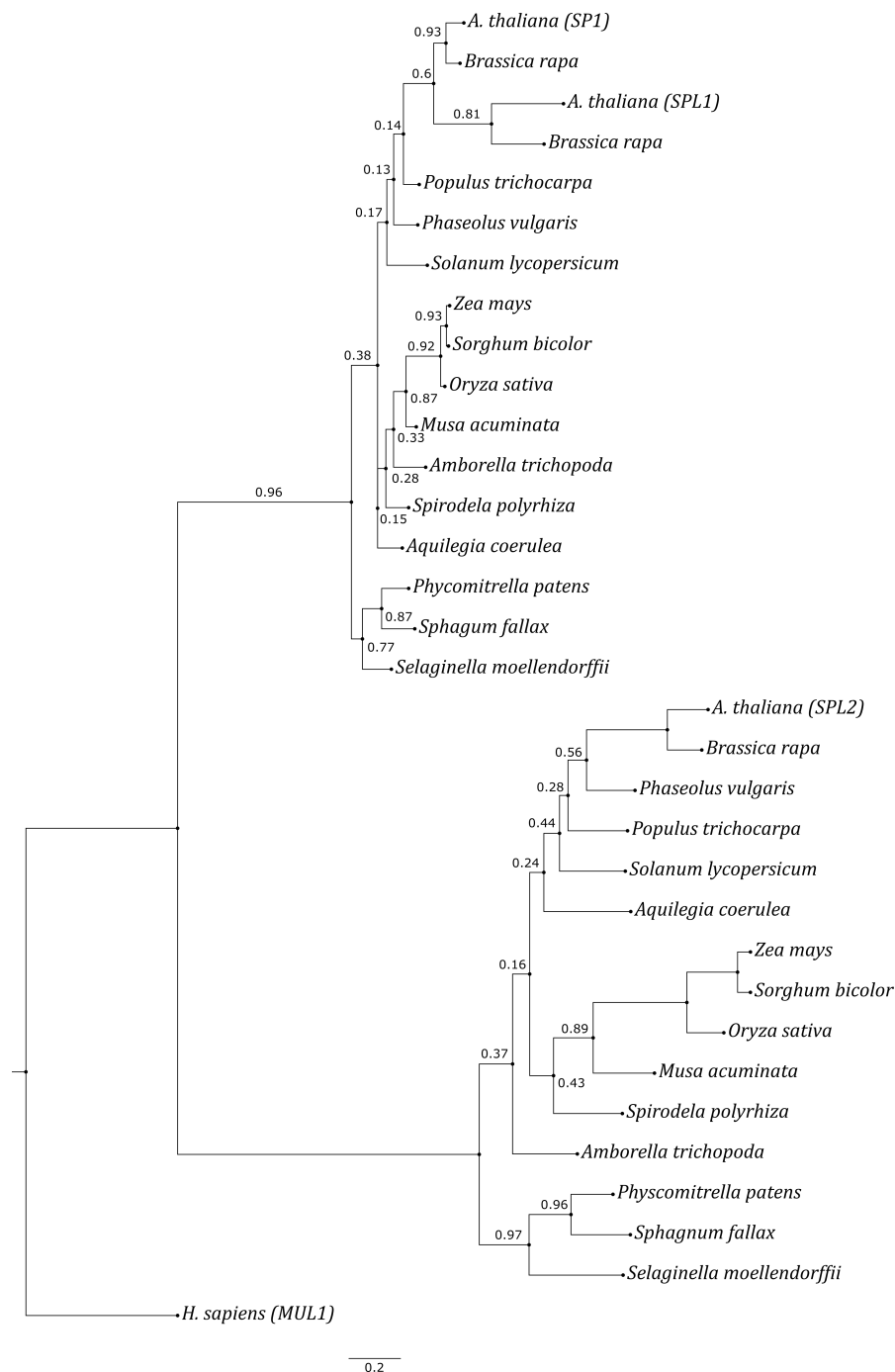


Figure 5.2. Boot-strapped maximum-likelihood phylogenetic tree of the SP1 family. Predicted protein sequences homologous to Arabidopsis SP1, SPL1 and SPL2, from a variety of plants, were retrieved from the Phytozome 12 database (Goodstein et al., 2012). The closest human relative to SP1, MUL1 is included as the outgroup. Sequences were aligned with MAFFT and the phylogeny was inferred using RAxML, and bootstrapped 100 times (Stamatakis, 2014). Only boot-strap values of < 1 are shown. Scale bar represents 0.2 changes per site.

Chapter 6

Materials and methods

6. Materials and methods

6.1 General use of *Arabidopsis*

6.1.1 Preparation of MS growth media

Seeds were typically germinated on half-strength Murashige and Skoog (MS) medium with 0.5% (w/v) sucrose and 0.8% phyto-agar (Duchefa, 950-1050 g/cm³). This was prepared in 4 L batches by mixing, 17.2 g of MS basal salt mixture (Duchefa), 20 g of sucrose (Fisher), and 2 g of 2-(N-morpholino) ethanesulfonic acid (MES [Melford]) in 3.7 L purified water (Elga Purelab Ultra); the pH was then adjusted to 5.7 with KOH, before adding more purified water to 4 L. The MS solution was added in 400 mL aliquots to 500 mL glass bottles, to which 3.2 g phyto-agar was added. The bottles were autoclaved at 111°C for 15 minutes, and stored at room temperature until further use. To prepare plates, an MS bottle was heated in a microwave on low power over 20-30 minutes, until completely melted. Liquid MS-phyto-agar was either directly poured into plastic Petri dishes (typically, Sterilin, 10 cm diameter, or 12 cm² plates were used) inside a laminar flow hood (Gelaire Flow Laboratories), or first supplemented with sterile antibiotics or salts, according to the experimental requirements. Poured plates (ca. 16 × [10 cm diameter plates] per 400 mL medium) were dried for at least one hour in the laminar flow hood and then stored at 4°C in plastic bags until further use.

6.1.2 Germination of *Arabidopsis* seeds on MS media

Arabidopsis thaliana was typically germinated on half-strength MS media, by stratification followed by transfer to a growth cabinet for 10-14 days, until they

were either sampled/processed for experimentation, and/or transferred onto soil, for growth until later stages of development.

Arabidopsis seeds were first surface-sterilised with ethanol. Seeds were aliquoted in microfuge tubes to an approximate volume of 20-50 μL . To each aliquot, 500 μL of 70% ethanol containing 0.05% Triton X-100 was added. The tubes were sealed and incubated horizontally at 250 revolutions per minute (rpm) on an orbital shaker (SO1 [Stuart Scientific]) for 10 minutes. Then they were briefly centrifuged in an Eppendorf centrifuge (5415 D, rotor EL 082) at low speed to collect the seeds at the bottom of the tube. The supernatant was removed and replaced with 500 μL of >99.8% ethanol. The tubes were again incubated on the orbital shaker for 10 minutes. For each aliquot of seeds, a Whatman filter paper (70 mm diameter) was folded in the middle and opened again to leave a groove for the seeds to be easily dispersed on the plates. The filter papers were sterilised by soaking with ethanol (>99.8% ethanol [Honeywell]) in a laminar flow hood and let dry. The seeds were centrifuged and transferred onto the dry filter paper with a 1 mL tip, cut 2-3 mm from the tip to widen the aperture. The seeds were left to dry and then dispersed on the plates to a density of roughly 100-150 seeds per plate. Plates were sealed with Micropore tape (3M) and stratified in the dark at 4°C for 2-3 days. Stratified plates were put into Percival growth cabinets with white fluorescent light of approximately 120 $\mu\text{E m}^{-2} \text{s}^{-1}$ and a constant temperature of 20°C with a long day cycle (8 hours dark, 16 hours light).

6.1.3 Growth of Arabidopsis plants on soil

Arabidopsis plants were transferred to soil at 10-14 days old (10-14 days post-transfer from stratification to growth cabinet). To propagate plants for seed,

for floral dipping, or for collection of tissue to make protoplasts, seeds were sometimes germinated directly on soil.

Soil was prepared, by first mixing a 4:1 ratio of compost (Levington M2) to vermiculite (Sinclair Pro, medium, 2-5 mm); then 2 g of pesticide (Exemptor®) per 10 L of the soil mix was added to prevent propagation of Sciarid fly. Soil was poured into pots of desired size. The soil was rehydrated by allowing pots to absorb water from the base of the pot over the course of one hour.

To germinate seeds on soil, seeds of the necessary genotype were scattered onto hydrated soil, ~10 seeds per pot. These were transferred directly to greenhouse. After two weeks some plants were removed, leaving 1-5 plants per pot (depending on pot size and requirement).

Alternatively, 10-14-day-old seedlings from plates were transferred onto soil using forceps. Care was taken so that roots remained intact; the roots were then completely covered by soil, to prevent the seedlings from dehydrating. The trays were put into growth cabinets (Percival) at 20°C, 60% relative humidity, with a long-day cycle and 120 $\mu\text{E m}^{-2} \text{ s}^{-1}$ of fluorescent white light for experiments, or in a glasshouse, with supplementary lighting to retain a long-day cycle.

All trays were covered with a plastic lid for 4-7 days to retain high humidity, then the lid was removed and the trays were watered regularly.

6.1.4 Salt-stress assay

For this experiment seeds which were harvested 16 months prior were used. MS media was prepared according to Section 6.1.1, with the addition 1% sucrose, and the addition of 150 mM NaCl. Seeds of the desired genotypes were

germinated on the media, as described in Section 6.1.2. The plants were left to grow in the plates for 14 days before assessment by photography, and chlorophyll quantification.

6.1.5 Crossing of Arabidopsis plants

Plants were grown on soil as described above. Desired genotypes were designated as the female acceptor and the male donor. Usually, homozygous mutants with a clear, recessive phenotype were used as female acceptors because self-pollinated progeny could be easily eliminated in the F1 based on appearance of the mutant phenotype. Mutants containing resistance markers were used as male donors because in this case selection could be used in the F1 to eliminate self-pollinated progeny. From 2-3 branches of the acceptor genotype per cross, all siliques and opened flowers were removed. From the 3-4 largest unopened flower buds of these branches, all sepals, petals and stamens were removed with fine-tipped forceps. The branches were labelled with the donor genotype and open donor flowers were used to pollinate the stigma using forceps.

6.1.6 Floral-dip transformation of Arabidopsis plants

Floral-dip transformation of Arabidopsis was adapted from a method described previously (Bent, 2006). Plants were grown on soil as described above. When the plants reached a height of approximately 5 cm, the primary shoots were cut in order to promote growth of secondary branches. The plants were left to grow secondary branches for approximately one week.

Agrobacterium tumefaciens colonies were selected on antibiotic selection LB plates, incubated at 28°C for 1-2 days. Single colonies were propagated in 5 mL of LB medium containing the required antibiotic. The culture was incubated

at 28°C overnight, shaking at 240 rpm. Then, 1 mL of the overnight culture was transferred into 300 mL LB medium containing the required antibiotic in a 1 L Erlenmeyer flask. The culture was incubated at 28°C overnight, shaking at 240 rpm.

Overnight *Agrobacterium* culture containing the necessary plant transformation vector (300 mL) was centrifuged at 5000 xg for 10 minutes in 500 mL Nalgene tubes using a Sorvall RC6 centrifuge (using the Swinging Bucket Rotor 7500 6445). The pellet was resuspended in approximately 50 mL of 5% (w/v) sucrose containing 300 µL/L Silwet L-77 (Lehle Seeds) and further diluted in sucrose and Silwet solution until the suspension reached an optical density (OD) between 0.7 and 0.8 at 600 nm wavelength. All siliques and open flowers of the Arabidopsis plants were removed and the above-ground parts were dipped into the *Agrobacterium* suspension with gentle swirling for 20-30 seconds. The plants were covered with a plastic dome for one day to retain high humidity and then watered as usual from ~5 days after transformation until full maturity. Mature seeds were sown on MS plates containing 200 µg/L cefotaxime (to prevent growth of *Agrobacterium*) and an additional selection agent for the specific T-DNA-borne plant resistance marker.

6.1.7 Chlorophyll measurements

Plants were grown on plates to an age of 10-14 days or on soil to an age of 28 days as described above. Amounts of chlorophyll in these plant samples were measured in two different ways.

6.1.7.1 Chlorophyll measurements of plants grown on MS

Samples of whole seedlings were weighed in microfuge tubes to a total of 15-30 mg depending on phenotype. To the seedlings 1 mL of dimethylformamide (DMF) was added and the tubes were wrapped in aluminium foil and incubated overnight on an orbital shaker at 4°C and 30 rpm. Wild type-like samples were diluted 1:5 in DMF before measurement. The absorbance of each sample was measured for the wavelengths 646.8 nm and 663.8 nm relative to a blank (DMF) using a spectrophotometer (Ultraspec 100 Pro Visible, Amersham) and a 200 µL quartz cuvette. The cuvette was rinsed with DMF between the reads. The readout of the absorbance should ideally be between 0.01 and 0.8; if it exceeded 0.8 the sample was further diluted in DMF. Amounts of chlorophyll in nmol per mg of fresh weight were calculated as follows (Porra et al., 1989):

$$\text{Chlorophyll } a = \frac{(13.43 \times A_{663.8}) - (3.47 \times A_{646.8})}{\text{fresh weight}} \times \text{dilution factor}$$

$$\text{Chlorophyll } b = \frac{(22.9 \times A_{646.8}) - (5.38 \times A_{663.8})}{\text{fresh weight}} \times \text{dilution factor}$$

$$\text{Chlorophyll } a + b = \frac{(19.43 \times A_{646.8}) + (8.05 \times A_{663.8})}{\text{fresh weight}} \times \text{dilution factor}$$

6.1.7.2 Chlorophyll measurements of plants grown on soil

Individual leaves per sample were measured using the SPAD-502 chlorophyll meter (Konica Minolta) according to the manufacturer's guidelines. The readout, x , was converted into nmol of total chlorophyll ($a + b$) per mg of fresh weight using the following formula (Ling et al., 2011):

$$\text{Chlorophyll } a + b = (0.0007x^2 + 0.023x) + 0.0544$$

6.1.8 Measurement of rosette diameter

The rosette diameter of 28-day-old soil grown plants was measured by recording the largest distance between the tips of two rosette leaves. One measurement was taken per plant.

6.1.9 TEM images

Seedlings were grown to 10 days old and processed at University of Leicester and were brought to the Electron Microscopy Laboratory, University of Leicester. Cotyledons were used for the analysis. All samples were processed and photographed by Natalie Allcock.

Images were analysed using Adobe Photoshop® software. Length and width were measured at the longest and widest points of the chloroplast cross-sections, respectively. The chloroplast cross-section plan area was calculated as:

$$\pi \frac{(length \times width)}{2}$$

6.2 General use of bacteria

6.2.1 Preparation of LB medium

Lysogeny broth (LB) was made from, 10 g of tryptone (Melford), 5 g of yeast extract (Melford), and 5 g of NaCl (Sigma), dissolved in 1 L of purified water. The broth was aliquoted into 500 mL Duran bottles and autoclaved for 15 minutes at 120°C. For LB agar plates, 1% by weight of agar (Melford) was added to the liquid medium prior to autoclaving. Before use the LB agar was melted in the microwave and left to cool to below 60°C, at this point any antibiotics required according to the experimental requirements were added. The medium was poured into plates in a flow hood and left to dry for at least one hour.

6.2.2 Heat-shock transformation of *E. coli* and plasmid isolation

Competent DH5α *Escherichia coli* cells (Alpha-Select Silver Efficiency [Bioline]), aliquoted to 25 µL were thawed on ice. About 10-50 ng (or according to purpose) of plasmid was added to the cells and the tubes were incubated on ice for 30 minutes with flicking every 5-10 minutes. The cells were heat-shocked for 45 seconds at 42°C in a water bath and then left on ice for 2 minutes. Liquid LB medium (450 µL), was added and the cells, which were then incubated at 37°C for 2 hours in a water bath. Aliquots (50 µL and 200 µL) of each transformation were streaked on LB plates containing a plasmid-specific antibiotic, and the plates were incubated at 37°C overnight. Individual colonies were scraped with a pipette tip, and the whole tips were put in 5 mL of liquid LB containing the plasmid-specific antibiotic. The culture was incubated at 37°C overnight with shaking at 240 rpm. *E. coli* stocks were prepared with a 1:1 mixture of overnight culture and sterile 50% glycerol and stored at -80°C. Minipreps were performed with the

overnight culture using the ISOLATE II Plasmid Mini kit (Bioline) following exactly the manufacturer's guidelines. For midi-preps, 500 µL of the overnight culture was diluted in 50 mL of liquid LB medium containing the plasmid-specific antibiotic. The 50 mL culture was grown at 37°C overnight and midi-preps were performed with the culture using the QIAfilter Plasmid Midi kit (Qiagen) following exactly the manufacturer's guidelines. Plasmid DNA was quantified using a NanoPhotometer® P330 (IMPLEN).

6.2.3 Blue-white screening

Blue-white screening was used to identify *E. coli* colonies transformed with pBluescript SK⁻, containing correctly cloned DNA. This was done by adding 40 µL 20 mg/mL 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside (X-Gal) dissolved in DMSO, and 40 µL 0.1 M isopropyl thiogalactoside (IPTG) to the bacteria suspension prior to spreading onto the LB agar plate. After transformed bacteria, plus X-Gal and IPTG were spread onto an LB agar plate containing 100 µg/mL ampicillin, for selection of pBluescript, the plates were incubated at 37°C overnight. After 16-24 hours colonies were visible. Colonies transformed with pBluescript SK⁻, containing the insert of interest would appear white, whereas those transformed with pBluescript SK⁻ which did not contain the insert, appeared blue. White colonies were therefore propagated.

6.2.4 Freeze-thaw transformation of *Agrobacterium*

This method was adapted from a method described previously (Weigel and Glazebrook, 2006). About 1 µg of the desired binary plasmid was added to 50 µL of *Agrobacterium tumefaciens* stock suspension (GV3101 cells) and mixed. The tube was frozen in liquid nitrogen and subsequently thawed by incubating at 37°C

in a water bath for 5 minutes. To the cells, 1 mL of LB medium (see Section 6.2.1) was added, and they were incubated at 28°C for 2-4 hours, shaking at 240 rpm in a temperature controlled cabinet. After 30 seconds centrifugation at 5000 xg (Eppendorf centrifuge, 5415 D, rotor EL 082) the pellet was resuspended in 50-100 µL of LB medium and spread on LB plates containing gentamycin (7.15 µg/mL), for selection of the Ti (tumour inducing) plasmid (pMP90RK), plus a specific antibiotic for selection of agrobacteria transformed with the desired binary vector. The plates were incubated at 28°C for 2-3 days. Then, single resistant colonies were propagated.

6.3 DNA related techniques

6.3.1 PCR, agarose gel electrophoresis, and gel band purification

Polymerase chain reaction (PCR) was performed using Biometra Thermocycler T1 or T3 devices and standard reaction mixtures (10 mM Tris-HCl [pH 8.3], 50 mM KCl, 50 mM MgCl, 0.05% Tween[®] 20, 200 μ M of each dNTP, 125 nM of each primer) with differences based on protocols for different DNA polymerases. DNA polymerases were mostly used at 0.5 units per 20 μ L reaction, or otherwise depending on the user guide. Primers were designed using Primer-BLAST (Ye et al., 2012), and/or PrimerQuest[®], and ordered from Sigma-Aldrich. Agarose gels were prepared with 1% (standard) to 3% (dCAPS) agarose (Melford) in 0.5 $\times$ TBE (44.5 mM Tris, 44.5 mM boric acid, 1 mM EDTA) containing 1 $\times$ SYBR Safe DNA gel stain (Invitrogen). Samples were loaded in 1 $\times$ Orange G loading buffer (0.1% Orange G, 50% glycerol), next to a DNA marker (1 Kb Plus DNA Ladder [Invitrogen]) and run in tanks (Mini-Sub[®] Cell GT, Bio-Rad) containing 0.5 $\times$ TBE at $\sim$ 10 V/cm. Finally, the bands were visualized using GeneFlash (Syngene) or LAS4000 (GE). If the PCR product was to be purified, the gel band was cut on a UV dual intensity transilluminator (UVP) at low intensity and processed using the GenElute[™] Gel Extraction kit (Sigma) following exactly the manufacturer's instructions.

6.3.2 Restriction digestion

Restriction enzymes (all from NEB) were used to cleave DNA at specific sites. Plasmid and genomic DNA sequences were first digested in silico, mostly using the online program NEBcutter (NEB). When multiple digestions were needed, NEB Double Digest Finder (NEB) was also used, in order to choose

compatible restriction enzymes, and optimal reaction conditions for certain experiments. Restriction digestions were performed using buffers and conditions recommended for the enzyme or combination of enzymes. Digested fragments were visualized on 1% (cut vectors and inserts), or 3% (dCAPS markers) agarose gels.

6.3.3 Cloning of a DNA sequence using the Gateway™ technology

Primers for the DNA sequence to be cloned were designed based on sequence information in The Arabidopsis Information Resource 10 (TAIR) database (Lamesch et al., 2012). The forward primer was given the 5'-addition; 5'-AAAAAGCAGGCTCC-3', and the reverse primer was given the 5'-addition; 5'-AGAAAGCTGGGT-3'. The PCR of DNA inserts for cloning was always performed using Phusion High-Fidelity DNA polymerase (Thermofisher). The PCR product was run on a 1% agarose gel, and was extracted, and then a second PCR was performed (annealing temperature 55°C) using 100 ng of the gel-extracted primary PCR product in a 50 µL reaction, using the primers: AttB1; 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCT-3', and AttB2; 5'-GGGGACCACTTTGTACAAGAAAGCTGGGT-3', which bind to the 5'-additions that were introduced into the primary PCR product with the first primers. The secondary PCR product was run on a standard 1% agarose gel and retrieved using the QIAquick Gel Extraction kit.

6.3.3.1 BP reaction

For the BP reaction, a piece of insert DNA flanked by attB recombination sites was transferred into pDONR201 by recombination. For this reaction 1:1 molar ratio of insert to vector were incubated with the BP clonase II (Invitrogen).

This was calculated as: $DNA\ required\ (ng) = 16.5 \times DNA\ length\ in\ kb$. The reaction was performed by adding the calculated amounts of insert and donor vector, with 1 μ L of BP clonase (Invitrogen) per 5 μ L reaction. The reaction was incubated at 25°C overnight and terminated by the addition of 0.5 μ L of proteinase K (Invitrogen) at 37°C for 10 minutes. *E. coli* DH5 α cells were transformed with the reaction as described in Section 6.2.2, and spread on LB plates containing the appropriate antibiotic to select for the donor vector. Plasmid from positive colonies was collected by mini-prep, and tested by a restriction digestion, as described in Section 6.3.2, so that the fragment pattern seen could distinguish between empty vectors and plasmids containing the desired insert. Plasmids containing the PCR product were sequenced using the pDONR-specific primers pDONR-F: 5'-TCGCGTTAACGCTAGCATGGATCTC-3' and pDONR-R: 5'-GTAACATCAGAGATTTTGAGACAC-3'. Only plasmids with the correct PCR insertion and sequence were used for LR reactions.

6.3.3.2 LR reaction

For the LR reaction, a pDONR201 vector containing a sequence of interest would donate that sequence to a desired destination vector by recombination. For this reaction, a 1:1 molar ratio of donor to destination vector was incubated with the LR clonase II (Invitrogen). This was calculated as: $DNA\ required\ (ng) = 16.5 \times DNA\ length\ in\ kb$. The reaction was performed by adding the calculated amounts of insert and donor vector, with 1 μ L of LR clonase (Invitrogen) per 5 μ L reaction. The reaction was incubated at 25°C overnight and terminated by the addition of 0.5 μ L of proteinase K (Invitrogen) at 37°C for 10 minutes. *E. coli* DH5 α cells were transformed with the reaction as described in Section 6.2.2 and spread on LB plates containing the destination vector-specific antibiotic. Mini-

preps were extracted from resistant clones and the plasmid was tested by restriction digestion as described in Sections 6.3.2. For stable expression of the constructs in *Arabidopsis*, *Agrobacterium tumefaciens* was transformed with the plasmid using the freeze-thaw method, and plants were transformed with *Agrobacterium* using the floral-dip method as described in Sections 6.2.4 and 6.1.6.

6.3.4 Cloning using digestion and ligation

In some cases, cloning was done by linearizing the vector with one (for blunt-end ligation) or two different restriction enzymes (sticky-end ligation), and then ligating the desired insert DNA between the digested sites; this was typically done in the vectors multiple cloning site (MCS). To generate an insert for blunt end ligation, the desired sequence was amplified with specific primers using a polymerase such as Phusion, which would not leave the product with overhangs. To generate an insert with sticky-ends, the primers used to amplify the desired piece of insert DNA would create restriction sites at each end of the PCR product, which when digested, would leave overhangs complementary to those of the digested vector. Digestions were carried out according to Section 6.3.2. Digested PCR products and vectors were gel extracted from 1% agarose gels, using GenElute gel extraction kit (Sigma). The ligation reaction was typically carried out by using a 3:1 molar ratio of insert to vector, and incubating overnight at 4°C with T4 ligase (Promega), according to manufacturer's protocol, in a 10 µL reaction volume. Subsequently 2 µL was used to transform 25 µL competent *E. coli* according to Section 6.2.2.

6.3.5 Constructs made

To produce *SP2OX*, the *SP2* CDS was amplified from cDNA using the primers: AttB1-SP2-F; 5'-
AAAAAGCAGGCTCCATGGGAGCTCAGAAGAGTATCCA-3', and AttB2-SP2-R; 5'-
AGAAAGCTGGGTTTGTTGATGAAGCAAGATTGGTG-3', containing the added sequences necessary for Gateway™ cloning (highlighted in red). The PCR product was cloned using the Gateway™ method, described in Section 6.3.3 into the destination vector pB2GW7, a binary plant overexpression vector from the Laboratory of Plant Systems Biology (PSB; Ghent University, Belgium).

To place *SP2* in pBluescript SK-, for in vitro translation, the *SP2* CDS was amplified from cDNA using the primers: KpnI-SP2-F; 5'-
GGTACCATGTCTTTAATGTTTCCTGCTTTCAGGG-3', and SP2-R; 5'-
CTATGTTGATGAAGCAAGATTGG-3'. The KpnI restriction site is highlighted in red. The product was gel purified and was cloned according to Section 6.3.4.

To produce a construct encoding *SP2-Myc*, the *SP2* CDS was amplified without the stop codon, using the primers: BclI-SP2-F; 5'-
GGGTAACTCT**GATCA**ATGGGAGCTCAGAAGAGTATCCA -3', and SP2-nostop-NotI; 5'-AAAC**CGG**TCTGCGGCCGCTGTTGATGAAGCAAGATTGGTG-3'. These make an *SP2* CDS with no stop codon, flanked with the BclI and NotI restriction sites (highlighted in red), so that PCR product could be digested and ligated directly into the modified Gateway™ entry vector, pE3c, which encodes a fusion of six repeats of the c-Myc epitope at the to the C-terminus of the inserted sequence (Dubin et al., 2008). This was then recombined into the binary vector pB2GW7 according to Section 6.3.3.2.

6.3.6 Plant DNA extraction and genotyping

One medium-sized leaf per sample was put into a 1.5 mL microfuge tube containing a 2.5-3.5 mm glass bead (BDH), and flash-frozen in liquid nitrogen. The leaf tissue was ground using a Qiagen TissueLyser for 30 seconds 25 Hz, and then 0.5 mL of DNA extraction buffer (200 mM Tris-HCl, [pH 7.5], 250 mM NaCl, 25 mM EDTA, 0.5% SDS) was added to the tissue powder. The samples were vortexed, then incubated on ice for minimum 5 minutes, and then centrifuged at maximum speed and 4°C for 5 minutes in an Eppendorf 5417 R microcentrifuge. The supernatant was transferred to a fresh tube and 450 µL of isopropanol was added to precipitate the DNA. The samples were incubated at -20°C for at least 1 hour and then centrifuged at maximum speed and 4°C for 20 minutes. The supernatant was discarded, the pellets were washed with 70% ethanol, and the samples were again centrifuged at maximum speed and 4°C for 5 minutes. The supernatant was discarded completely and the pellets were dried in a fume-cabinet for 15-20 minutes. Dried DNA pellets were resuspended in 50 µL of sterile water, by incubating for 2 hours to overnight, at 4°C.

6.3.6.1 Genotyping by dCAPS

The alleles genotyped by this method, as well as the primer sequences and necessary restriction enzymes, can be found in Table 6.2.

In cases where a single nucleotide polymorphism needed to be detected (such as the single base substitutions caused by EMS mutagenesis), a primer was designed to create a restriction site at the polymorphic locus (this could either make the wild type or mutant sequence the restriction site), this was done by introducing a sequence alteration in the primer that led to a full recognition site

in the PCR product of one genotype but not in the other. This derived cleaved amplified polymorphic sequence (dCAPS) primer was made in a way such that at least 20 bp could be cleaved off by the restriction enzyme, and together with the reverse primer, was designed so that the total length of the PCR product was around 100 bp (Neff et al., 1998). Treatment with the necessary restriction enzyme would lead to bands of approximately 100 bp for the non-cleaved product, and 80 bp for the cleaved product, which could be resolved on a 3% agarose gel to distinguish, homozygous and heterozygous genotypes.

6.3.6.2 Genotyping for T-DNA insertions

To correctly genotype a T-DNA insertion, two PCR reactions were performed. The first used a primer specific for the left-border of the T-DNA, and a complementary primer which was placed approximately 500 bp downstream of the putative T-DNA insertion site on the side closest to the left-border. This would confirm presence of the T-DNA in the genetic locus of interest. The second PCR used primers specific to the genetic locus of interest, which would amplify the region expected to contain the T-DNA. Amplification of this product would not occur under the PCR conditions used if a T-DNA were present, and in the absence of a T-DNA, the product of expected size for that genetic locus would be amplified. Amplification in both reactions indicated heterozygosity, whereas successful amplification in only one reaction indicated homozygosity.

For *sp2-4*, the primers used to identify the T-DNA in the *SP2* locus were: Lba1; 5'-TGGTTCACGTAGTGGGCCATCG-3', and SP2-R3; 5'-TACAGCTTCTCCTTGGACTGT-3'. The primers to amplify the *SP2* locus containing the T-DNA insertion site were: SP2-F3; 5'-

CGTGGAGCTGTTGCATGAT-3', and SP2-R3; 5'-TACAGCTTCTCCTTGGACTGT-3'.

For *ppil-2* the primers used to identify the T-DNA in the *PPII* locus were: 127ex2F; 5'-GGTCTCTCGTTTCGTGAATGG-3', and Lba1; 5'-TGGTTCACGTAGTGGGCCATCG-3', and the primers to amplify the SP2 locus containing the T-DNA insertion site were: 127ex2F; 5'-GGTCTCTCGTTTCGTGAATGG-3', and 127ex3R; 5'-AAAGGACTGACACGGACGAC-3'.

6.3.7 Verification of *sp2* point mutants

The EMS derived *sp2* point mutations identified by whole genome sequencing (Section 2.2.1) were verified by Sanger sequencing of *SP2* cDNA from each *sp2* allele. The cDNA was amplified with the primers: 5'-GGTTGTGTCCAAGTGGCTTA-3', and 5'-TACAGCTTCTCCTTGGACTGT-3'. The products, including both products observed for *sp2-1*, were gel purified and Sanger sequenced using the primer, 5'-GGTTGTGTCCAAGTGGCTTA-3'.

6.4 RNA-related techniques

6.4.1 RNA extraction and cDNA synthesis

Arabidopsis seedlings were grown on plates for two weeks (Section 6.1.2). For each sample, 100 mg was harvested into 1.5 mL microfuge tubes each containing two 2.5-3.5 mm glass beads (BDH), and flash-frozen in liquid nitrogen. The tissue was ground using a Qiagen TissueLyser for 60 seconds 25 Hz. A Spectrum™ Plant Total RNA Kit (Sigma) was used for RNA extraction following exactly the manufacturer's guidelines, with the addition of the optional on-column DNase treatment (On-Column DNase I Digestion Set [Sigma]) to remove DNA. The isolated RNA was then quantified using a NanoPhotometer® P330 (Implen).

For reverse transcription, the Superscript® IV system (Invitrogen) was used, according to manufacturer protocol. Reactions were prepared to a 20 µL total volume, and per reaction, 5 µg of RNA was used. Post cDNA synthesis, samples were diluted 1:10.

6.4.2 Semi-quantitative RT-PCR

Semi-quantitative RT-PCR was performed using primers for the target gene, and the control gene primers (either *eIF4E*, or *Actin2*; Table 6.3). An equal amount of cDNA was used for each sample, and a low number of PCR cycles (20-25 cycles) was used to ensure PCR product would be doubling each cycle. The bands of the gel image were then quantified in silico using Aida Image Analyzer software (Raytest) by drawing boxes of equal size around each band, to quantify pixel values. The values for the target gene were normalized to the values for the

control bands. To make all results relative to wild type, each sample value was then divided by the average wild type value.

6.4.3 qPCR

For quantitative real-time PCR (qPCR), cDNA was produced as described above (Section 6.4.1), and primers specific for the target that produce a 80-120 bp product, and anneal at 60°C, were designed. Reactions were prepared in a 20 µL volume, using the qPCRBIO SyGreen Mix Hi-ROX (PCR Biosystems) reaction system in 96-well MicroAmp® PCR plates (Applied Biosystems), and run and analysed with a StepOnePlus™ (Applied Biosystems). For each genotype tested, four biological replicates, with three technical replicates of each, were run. Additionally, no-template controls were used for each amplicon. The expression level of the target in each sample was determined using the comparative C_t method in StepOne™ (Applied Biosystems) software, and normalized to both *GAPDH* and *Actin2*.

6.5 Protein analysis

6.5.1 Total protein extraction from Arabidopsis plants

Plant material (leaf or seedlings; 50 mg) was harvested into 1.5 mL microfuge tubes each containing two 2.5-3.5 mm glass beads (BDH), and flash-frozen in liquid nitrogen. The tissue was ground using a Qiagen TissueLyser for 60 seconds 25 Hz. A protein extraction buffer (EB) was prepared (100 mM Tris-HCl, [pH 6.8], 10% glycerol, 0.5% SDS, 0.1% Triton X-100, 5 mM EDTA, 10 mM DTT, 1 × cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail [Roche]); the DTT and plant protease inhibitors were added immediately before use. Per 50 mg of tissue, 300 µL of EB was added, the samples were then vortexed and incubated on ice for 30 minutes, with intermittent vortexing. The samples were subsequently centrifuged at maximum speed and 4°C for 10 minutes in a microcentrifuge. About 250 µL of the supernatant per sample was transferred to a fresh tube and stored at -20°C (short term) or -80°C (long term), or used immediately for downstream experiments.

6.5.2 Protein quantification by Bradford assay

Bovine serum albumin (BSA [NEB]; 10 µg/µL) was diluted 1:10 in sterile water to 1 µg/µL. Standards were made by adding 0, 1, 2, 4, 6, 8 and 10 µL of 1 µg/µL BSA to 7 microfuge tubes containing sterile water, to reach a total volume of 799 µL each. Then, 1 µL of protein extraction buffer was added to compensate for its absorption. For the actual samples, 1 µL of protein (in extraction buffer) was added to 799 µL of sterile water. Afterwards 200 µL of Bio-Rad Protein Assay (Bio-Rad) was added to each tube, and the samples were vortexed and incubated for 5 minutes at room temperature. The absorbances of the samples

were then measured at a wavelength of 595 nm using NanoPhotometer® P330 (Implen), and disposable cuvettes (Semimicro [Fisherbrand]). The control absorbance values were plotted against the BSA concentrations automatically in the onboard software of the NanoPhotometer®, and the equation of a linear approximation was used to calculate the concentrations of all other samples.

6.5.3 Preparation of SDS-PAGE gels

For SDS-PAGE, either precast 4–15% Mini-PROTEAN® TGX™ gels (Bio-Rad), or gels were prepared by hand were used. Polyacrylamide resolving gels containing 375 mM Tris-HCl, (pH 8.8) and 0.1% SDS were generally prepared with acrylamide concentrations between 10% (w/v; lower resolution, shorter running time) and 15% (higher resolution, longer running time). A 30% (w/v) acrylamide stock containing 0.8% (w/v) bis-acrylamide (ProtoGel) was diluted in the Tris-SDS buffer to reach the required concentrations. The stacking gel always contained 125mM Tris-HCl, (pH 6.8), 0.1% SDS and 5% (w/v) acrylamide. Before pouring, 50 µL of 10% (w/v) ammonium persulfate (AP) and 5 µL tetramethylethylenediamine (TEMED) were added in order to polymerize the acrylamide. To pour and run the gels, the Mini-PROTEAN 3 Cell gel system (Bio-Rad) was used. To make the resolving gel 3.33 mL of the liquid preparation was added to the mould, and immediately isobutanol was added to the top to maintain a flat surface and aid polymerization by excluding air. These were left for 45 minutes, and after polymerization of the resolving gel, the isopropanol was decanted and remaining droplets were removed with a filter paper. The stacking gel was then poured on top, and a comb was inserted to form the wells. These were left for 45-90 minutes to set.

6.5.4 SDS-PAGE

Then the comb was and any packaging was removed from the desired polyacrylamide gels, and then the wells were washed with distilled water to rinse away any debris. Polyacrylamide gels were fitted into the Mini-PROTEAN gel chamber which was then filled with enough 1 × TGS buffer (25 mM Tris, 192 mM glycine, 0.1% SDS) to cover both electrodes. Prior to sample loading, all protein samples were diluted in an equal volume of 2 × Laemmli buffer (4% SDS, 20% glycerol, 120 mM Tris-HCl [pH 6.8], 50 mM DTT and 0.02% bromophenol blue; Laemmli, 1970), and the samples were heated to 95°C for 5 minutes in order to denature the proteins. Denatured protein samples were loaded into the wells of the gel alongside 3 µL of a protein ladder (Precision Plus Protein™ Dual Color Standards [Bio-Rad]) as a standard. Gels were run at 180 V using a Power Pac 300 station (Bio-Rad) until the bromophenol blue dye front reached the bottom of the resolving gel. Then, gels were either directly used for western blotting or Coomassie staining.

6.5.5 Coomassie staining

Gels were first fixed for 30 minutes in 40% methanol 10% acetic acid solution, then incubated in Coomassie stain (0.25% Coomassie Brilliant Blue R-250, 45% methanol, 10% acetic acid), for 2 hours to overnight. To remove the dye from the background, gels were de-stained with 40% methanol 10% acetic acid solution 3 times, up to overnight, until the background was clear.

6.5.6 Western blotting

Western blotting began with SDS-PAGE of desired protein samples as described in Section 6.5.4. Proteins separated in polyacrylamide gel were

transferred to a polyvinylidene fluoride (PVDF) membrane, Hybond P 0.45 PVDF (Amersham). This was done by first wetting the PVDF with methanol, then pressing the gel and membrane together in a cassette, with all items submerged in transfer buffer (25 mM tris, 192 mM glycine, 20% methanol); this was made as a sandwich: sponge, blotting paper, gel, PVDF, blotting paper, sponge. The cassette was closed and then placed in the Mini Trans-Blot Cell system (Bio-Rad), with the gel-side facing the cathode, and PVDF-side facing the anode. Transfer buffer was added to fill the tank, and transfer of proteins to the PVDF membrane was performed at 100 V for 60 minutes.

After transfer, the blot was cut into sections according to the sizes of proteins to be analysed, and was done with the help of the pre-stained protein ladder. The membranes were then blocked for 30 minutes, in TBS-T (10 mM Tris-HCl [pH 8.0] 150 mM saline 1% Tween® 20 [Sigma]), containing either 5% (w/v) milk powder (Marvel original) for most antibodies, or 1% (w/v) BSA (Sigma) for antibodies against Toc159, and Toc34. The membranes were then incubated overnight at 4°C with primary antibodies diluted in TBS-T, containing either 5% milk powder or 1% BSA, according to primary antibody. The membranes were then washed three times for 10 minutes with TBS-T, and subsequently incubated for 1 hour at room temperature in secondary antibody, which was diluted in TBS-T with 5% milk powder or 1% BSA, according to which had been use for the primary antibody. Secondary antibodies were used according to the organism in which the primary antibody had been raised, and were conjugated to horseradish peroxidase (HRP). Before detection, the membranes were again washed three times or 10 minutes with TBS-T. Western blots were visualized, using the EZ-ECL Chemiluminescence Detection Kit (Geneflow), following the manufacturer's

guidelines. The chemiluminescence was visualized and recorded using a LAS-4000 image analyzer (GE).

6.5.7 Protein semi-quantification

The bands of a western blot image were quantified in silico using Aida Image Analyzer software (Raytest) by drawing boxes of equal size around each band, to quantify pixel values. The values for the protein of interest were normalised to the values for the control protein. The blot of the protein of interest and control protein always came from the same sample and membrane to which proteins had been transferred. To make all results relative to wild type, each sample value was then divided by the average wild type value.

6.6 Experiments using isolated chloroplasts

Chloroplast isolations were performed according to Kubis et al., (2008); detergent-free apparatus was used, and the materials and solutions used were kept at 4°C for the duration of the experiment.

6.6.1 In vitro protein import into chloroplasts

In vitro protein import was performed according to (Aronsson and Jarvis, 2011). The radiolabelled substrate protein was synthesized by in vitro translation using the TNT® T7 Quick for PCR DNA (Promega) system. To do this, the coding sequence of interest was cloned into pBluescript, as this contains a T7 promoter. PCR was performed using Phusion high fidelity polymerase (Thermofisher) to amplify the T7 promoter together with the coding sequence of interest from the vector, this was done using the primers: M13-F; 5'-CGCCAGGGTTTTCCCAGTCACGAC-3', and M13-R; 5'-AGCGGATAACAATTTCACACAGGA-3'. The reaction was prepared using methionine containing S³⁵ (Perkin Elmer) as the radioactive label, and was performed according to the TNT® T7 Quick for PCR DNA (Promega) protocol.

The radiolabelled substrate was incubated with isolated chloroplasts in the presence of ATP, and samples of 150 µL containing 10⁷ chloroplasts were taken at 2.5 minute intervals over a 10-minute period, and processed according to (Aronsson and Jarvis, 2011). Import rate was measured by the amount of pre-RUBISCO processed to mature RUBISCO, visualised by SDS-PAGE and phosphor-imaging with Storm 860 (GE), which was then quantified by measuring pixel values using Aida software (Raytest).

6.6.2 Alkaline treatment of chloroplasts

Alkaline treatment to assess membrane localization and integration of chloroplast proteins was performed on isolated chloroplasts, based on previously reported methods (Chu and Li, 2011). In these experiments, 1×10^7 chloroplasts were used per sample for the alkaline treatment. For analysis, post-treatment the samples were solubilized in 60 μ L of $2 \times$ Laemmli buffer. Proteins were analysed by SDS-PAGE and western blotting.

6.6.3 Chloroplast protease protection assays

Protease protection assays were conducted on isolated chloroplasts, based on previously reported methods (Froehlich, 2011). In the case of these experiments, 2×10^7 chloroplasts were used per sample, plus two different concentrations of each protease were used. The experiments presented in this thesis additionally included controls, whereby prior to protease treatment, chloroplasts were resuspended in buffer containing 1% Triton X-100 (Thermofisher). This was to solubilize the chloroplast membranes to release the protein and so assess whether the proteases used were capable of digesting the analysed proteins. To retrieve the protein from these controls, 1 mL of 100% acetone chilled to -20°C was added to the samples to precipitate the protein, the precipitate was then centrifuged at 20,000 xg (Eppendorf 5417 R) for 30 minutes, and the supernatant was disposed of. The pellet was centrifuged again briefly at 20,000 xg so that residual supernatant could be withdrawn. To permit resolubilization, the pellet was not allowed to dry completely. All samples were resuspended in 60 μ L of $2 \times$ Laemmli buffer for analysis. Proteins were analysed by SDS-PAGE and western blot.

6.7 Methods for protein interaction experiments

6.7.1 Transient expression in Arabidopsis protoplasts

For some co-immunoprecipitation experiments, proteins of interest with epitope tags were generated in vivo by transient expression in Arabidopsis mesophyll protoplasts (TEAMP); this was based on the TAPE method of protoplast isolation for TEAMP (Wu et al., 2009). The method was scaled up, so that there would be enough for 1×10^6 protoplasts per sample. The protoplasts were transfected with 50 μ g of plasmid DNA per 1×10^6 protoplasts. The plasmids used were p2GW7 (PSB; Ghent University, Belgium) containing SP1-HA, or HA-CDC48, and p2GWY7 (PSB; Ghent University, Belgium) containing YFP-HA, and mock transfection with no plasmid DNA was conducted as a form of control. Transfected protoplasts (1×10^6) were incubated in 3 mL of culture buffer (500 mM Mannitol, 4 mM MES, 10 mM KCl, [pH 5.6]) at 20°C in the dark for 16 hours to allow protein expression.

6.7.2 Co-immunoprecipitation

The entire procedure was conducted on ice, unless otherwise state. Isolated chloroplasts (1×10^8 per sample) or transfected protoplasts (1×10^6 per sample) were recovered by centrifugation, (chloroplasts 1000 xg, protoplasts 100 xg [Eppendorf 5417 R]) for 2 minutes. As much supernatant was removed as possible. Chloroplasts were lysed by the addition of 3 mL of hypotonic lysis buffer (25 mM HEPES [pH 8.0], 4 mM MgCl_2). The pellet was broken up with a sterile plastic pestle and the sample was then rotated in the cold room for 1 hour to allow efficient lysis. The membranes were then recovered by ultra-centrifugation (Sorvall Discovery M120 SE), at 110,000 xg (S100 AT6-0334 rotor) for 1 hour.

To the pelleted chloroplast membranes, or protoplasts, 1.5 mL of co-immunoprecipitation (co-IP) buffer (25mM Tris-HCl [pH7.5], 150mM NaCl, 10% glycerol, 1 % Triton X-100, 1mM EDTA, 1 × cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail [Roche]) was added, and a sterile plastic pestle was then used to break up the pellet. The sample was then rotated for 30 minutes at 4°C to allow efficient solubilisation. To recover debris, the sample was centrifuged at 10,000 xg for 20 minutes at 4°C; the supernatant was recovered and the pellet discarded.

For the chloroplast membrane samples, the protein content was quantified. The samples were then diluted equal to the sample with the lowest concentration. Typically, 1-1.2 mg/mL. An equal amount of the final protein preparation was used from each sample for the experiment.

In this thesis SP2-Myc was used as bait in these experiments, and EZview™ Red Anti-c-Myc beads (Sigma) were used for co-IP. To 1.5 mL microfuge tubes, 20 µL of the resuspended 50% anti-c-Myc bead slurry was added by a cut 200 µL pipette tip. The aliquoted beads were washed three times with 500 µL co-IP buffer; to do this, the co-IP buffer was added to the beads, and the tubes were mixed, left on ice for 1 minute, then centrifuged at 1000 xg for 1 minute at 4°C, and the supernatant was discarded. After washing, 1.4 mL of the solubilized protein was added to the beads, and then the tubes were rotated for 4 hours to overnight at 4°C.

The beads were then recovered by centrifugation at 1000 xg for 1 minute at 4°C. The flow-through was withdrawn, and the beads were washed again three times with 500 µL co-IP buffer, as described previously. Elution was performed by the addition of 30 µL of 10 µg/µL c-Myc peptide, followed by incubation on ice

for 10 minutes. The eluate was then collected by centrifugation at 1000 xg for 1 minute at 4°C. The samples taken throughout the experiment were then analysed by SDS-PAGE.

6.7.3 In vitro pulldown

Recombinant expression and purification of GST and GST-SP1flex, using glutathione agarose beads (Sigma) was completed as described previously (Ling et al., 2012). Radiolabelled SP2 was translated in vitro, as described in Section 6.6.1.

To prepare the beads for in vitro binding, empty glutathione agarose, GST and GST-SP1flex beads (30 µL) were washed once with 250 µL binding buffer (25 mM Tris-HCl [pH7.5], 5 mM MgCl₂, 1 mM DTT, 150 mM NaCl, 0.05 mM ZnCl₂, 0.1% Triton X-100, 10% glycerol). The beads were recovered by centrifugation at 1000 xg for 30 seconds, and then the supernatant was discarded. To prevent non-specific binding of the protein of interest, the beads were then blocked by the addition of 250 µL blotting buffer (binding buffer containing 5% [w/v] BSA). The beads were rotated with the blotting buffer overnight at 4°C. Before in vitro binding and pulldown, the beads were washed twice again with binding buffer as previously described.

To control for non-specific binding of radiolabelled SP2 to glutathione agarose blocked with BSA, the in vitro translated product was incubated with the blocked empty glutathione agarose beads, prior to incubation with the GST or GST-SP1flex beads. To these blocked empty glutathione beads, 750 µL binding buffer was added, followed by 30 µL of in vitro translated product. The beads and in vitro translated product were then rotated at 4°C for 45 minutes. The mixture

was then centrifuged for 10 min at 10,000 xg at 4°C, and then the supernatant was recovered, while the beads were discarded.

For in vitro binding, 260 µL of the supernatant was added to the GST and GST-SP1flex beads. These were then rotated at 4°C for 2 hours. The beads were recovered by centrifugation at 10,000 xg for 1 minute, and the supernatant was then discarded.

The beads were then washed six times with 250 µL wash buffer (25 mM Tris-HCl [PH7.5], 5 mM MgCl₂, 1 mM DTT, 500 mM NaCl, 0.05 mM ZnCl₂, 0.1% Triton X-100, 10% glycerol), by rotating for 2-3 minutes, followed by centrifugation at 10,000 xg for 1 minute. After the last wash, GST and GST fusion proteins were eluted from the beads by incubating with 50 µL elution buffer (200 mM Tris-HCl [PH8.0], 150 mM NaCl, 0.1% Triton X-100, 50 mM reduced glutathione [Melford]), at RT for 15 minutes with occasional mixing by flicking. The beads were then centrifuged at 10,000 xg for 1 minute. The supernatant was recovered and transferred to 50 µL Laemmli buffer.

The experiment was analysed by SDS-PAGE stained with Coomassie to visualize the GST fusion proteins, and then the same Coomassie gel was dried and analysed by phosphor-imaging by Storm 860 (GE) to visualize the radiolabelled protein.

6.8 Phylogenetic analysis

Sequences were obtained by BLAST searches of the Phytozome 12 database (Goodstein et al., 2012). Sequences were aligned by multiple alignment using fast Fourier transform (MAFFT; Katoh and Standley, 2013), and manual alignment adjustments were made using Mesquite 1.12 (Tangient). Phylogenies inferred by maximum likelihood used RAxML, and were tested by bootstrapping (Stamatakis, 2014).

For the phylogeny inferred using MrBayes 3.2 software (Ronquist et al., 2012), two runs were performed, each using 8 MCMC chains for 2 million generations, with temperature set to 0.2. Any parameter not mentioned remained at the default value. Trees were sampled every 1000 generations, totalling 8000 trees. Burn-in was set to 25%, and so the first 2000 were discarded. The resulting phylogeny was a minimum 50% consensus of the remaining 6000 sampled trees.

6.9 Tables

Mutant allele	Nature of mutation	Genotyping method
<i>sp2-1</i>	Point mutation	Table 6.2
<i>sp2-2</i>	Point mutation	Table 6.2
<i>sp2-3</i>	Point mutation	Table 6.2
<i>sp2-4</i>	T-DNA insertion	Section 6.3.6.2
<i>sp2-5</i>	Point mutation	Sanger sequencing
<i>sp1-1</i>	Point mutation	Table 6.2; (Ling et al., 2012)
<i>ppi1-1</i>	T-DNA insertion	(Jarvis et al., 1998)
<i>ppi1-2</i>	T-DNA insertion	Section 6.3.6.2
<i>ppi2-3 (fts1)</i>	Point mutation	(Woodson et al., 2015)
<i>toc75-III-3</i>	Point mutation	Table 6.2; (Stanga et al., 2009)
<i>tic22-III-1</i>	T-DNA insertion	(Kasmati et al., 2013)
<i>tic22-IV-2</i>	T-DNA insertion	(Kasmati et al., 2013)
<i>tic40-4</i>	T-DNA insertion	(Kovacheva et al., 2005)
<i>hsp93-V-1</i>	T-DNA insertion	(Kovacheva et al., 2005)

Table 6.1. Arabidopsis mutants used in this study.

Primer Name	Mutant	Sequence: 5' to 3'	Notes
SP1-dCAPS-F	<i>sp1-1</i>	CTCCAATATCGTCCTTGA A AGC	Enzyme: HindIII
SP1-dCAPS-R		TGTTGATAGAGATACATAGCAC	Digests: <i>sp1-1</i>
SP2-dC1-F	<i>sp2-1</i>	TGTTGCGGAATTGGTTTCAT T	Enzyme: DdeI
SP2-dC1-R		GGTTGTGTCCAAGTGGCTTA	Digests: WT
SP2-dC2-F	<i>sp2-2</i>	CATTGGTGGGCTAGGCAGTG ATC	Enzyme: BclI
SP2-dC2-R		CATGCAACAGCTCCACGTACCAA	Digests: <i>sp2-2</i>
SP2-dC3-F	<i>sp2-3</i>	CCACAATCCAAGAAGATGGT AC	Enzyme: KpnI
SP2-dC3-R		GCCTTGTTGCCAATACAGAA	Digests: WT
Toc75-III-3-dC-F	<i>toc75-III-3</i>	GCAAATCACAGGTGGAT CTG	Enzyme: DdeI
Toc75-III-3-dC-R		TGCCCATGGAGGACTAGAAC	Digests: <i>toc75-III-3</i>

Table 6.2. Primers used for genotyping by dCAPS. Genotyping using these primers was performed according to Section 6.3.6.1. The nucleotides in red are different from the target sequence, and are used to introduce the restriction site into the PCR product.

Primer Name	Gene	Sequence: 5' to 3'
Actin2-F	<i>Actin2</i>	TCAGATGCCCAGAAGTCTTGTTCC
Actin2-R	<i>AT3G18780</i>	CCGTACAGATCCTTCCTGATATCC
eIF4E1-F	<i>eIF4E1</i>	AAACAATGGCGGTAGAAGACACTC
eIF4E1-R	<i>AT4G18040</i>	AAGATTTGAGAGGTTTCAAGCGGTGTAAG
SP1-F1	<i>SP1</i>	GGTACAAGATAGTGCGTTGATG
SP1-R1	<i>AT1G63900</i>	GCTGTCAAGGACGATATTGGAG
Sp2-F2	<i>SP2</i>	GGTTGTGTCCAAGTGGCTTA
SP2-R2	<i>ATt3G44160</i>	CTATGTTGATGAAGCAAGATTGGTG
P36-F1	<i>P36 AT3G48620</i>	CAAGATGACAGAAGGGACC
P36-R1	(Primer pair 1)	GCCCATACCCGAATCCATAC
P36-F2		ATGTTTCATGGGATAAAGGGTT
P36-R2	(Primer pair 2)	CTATGTTGATGAAGCAAGATTGGT
P36-F3		GGGATAAAGGGTTGTATGATTCCAATG
P36-R3	(Primer pair 3)	GACCTTGGTCAGTAGCCTTT
P36-F4		CCGGACATTACCTTCTTCAGACA
P36-R4	(Primer pair 4)	GGCTTCCCTTGTCTCAGC
P36-F5		CATGCAGAATTCTCTGCTATGCA
P36-R5	(Primer pair 5)	TCCCTTCTGTCATCTTGTTCCG

Table 6.3. Primers used for RT-PCR experiments.

Primer Name	Gene	Sequence: 5' to 3'
Actin2-F	<i>Actin2</i>	TCAGATGCCCAGAAGTCTTGTTCC
Actin2-R	<i>AT3G18780</i>	CCGTACAGATCCTTCCTGATATCC
GAPDH-F	<i>GAPDH-B</i>	GTGGTGGATTTGGCTCACCT
GAPDH-R	<i>AT1G42970</i>	CTCATCAGCCGGGTTTGTCT
Toc159-QRT-F	<i>TOC159</i>	AACCTCTGAAGTGGCTAATAAGT
Toc159-QRT-R	<i>AT4G02510</i>	ACAACCTCTGGCTCTACA
Toc75-III-QRT-Fa	<i>TOC75-III</i>	TCGCATCTCCACTCAATC
Toc75-III-QRT-Ra	<i>AT3G46740</i>	GTCTCTGTATCTCGGTTAGG
Toc33-QRT-Fa	<i>TOC33</i>	AATGGTGAAGCGTGGATC
Toc33-QRT-Ra	<i>AT1G02280</i>	TGCTCCTTGAATCATCTTAACG

Table 6.4. Primers used for quantitative real-time PCR.

7. Bibliography

- Aronsson H, Jarvis RP** (2011) Rapid isolation of Arabidopsis chloroplasts and their use for in vitro protein import assays. *Methods Mol. Biol. Humana Press*, Clifton, N.J., pp 281–305
- Asada K** (2006) Production and scavenging of reactive oxygen species in chloroplasts and their functions. *Plant Physiol* **141**: 391–396
- Avci D, Lemberg MK** (2015) Clipping or extracting: Two ways to membrane protein degradation. *Trends Cell Biol* **25**: 611–622
- Baldrige RD, Rapoport TA** (2016) Autoubiquitination of the Hrd1 ligase triggers protein retrotranslocation in ERAD. *Cell* **166**: 394–407
- Baldwin A, Wardle A, Patel R, Dudley P, Park SK, Twell D, Inoue K, Jarvis P** (2005) A molecular-genetic study of the Arabidopsis Toc75 gene family. *Plant Physiol* **138**: 715–733
- Barsan C, Zouine M, Maza E, Bian W, Egea I, Rossignol M, Bouyssie D, Pichereaux C, Purgatto E, Bouzayen M, et al** (2012) Proteomic analysis of chloroplast-to-chromoplast transition in tomato reveals metabolic shifts coupled with disrupted thylakoid biogenesis machinery and elevated energy-production components. *Plant Physiol* **160**: 708–725
- Bauer J, Chen K, Hiltbunner A, Wehrli E, Eugster M, Schnell D, Kessler F** (2000) The major protein import receptor of plastids is essential for chloroplast biogenesis. *Nature* **403**: 203–207
- Bédard J, Kubis S, Bimanadham S, Jarvis P** (2007) Functional similarity between the chloroplast translocon component, Tic40, and the human co-chaperone, Hsp70-interacting protein (Hip). *J Biol Chem* **282**: 21404–21414
- Bent A** (2006) Arabidopsis thaliana floral dip transformation method. *Methods Mol Biol* **343**: 87–103
- Berardini TZ, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, Huala E** (2015) The arabidopsis information resource: Making and mining the “gold standard” annotated reference plant genome. *Genesis* **53**: 474–485

- Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, et al** (2014) SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* **42**: W252–W258
- de Bie P, Ciechanover A** (2011) Ubiquitination of E3 ligases: self-regulation of the ubiquitin system via proteolytic and non-proteolytic mechanisms. *Cell Death Differ* **18**: 1393–1402
- Blancaflor EB** (2013) Regulation of plant gravity sensing and signaling by the actin cytoskeleton. *Am J Bot* **100**: 143–152
- Bolte K, Gruenheit N, Felsner G, Sommer MS, Maier UG, Hempel F** (2011) Making new out of old: Recycling and modification of an ancient protein translocation system during eukaryotic evolution. *BioEssays* **33**: 368–376
- Bracher A, Whitney SM, Hartl FU, Hayer-Hartl M** (2017) Biogenesis and metabolic maintenance of Rubisco. *Annu Rev Plant Biol* **68**: 29–60
- Bragoszewski P, Gornicka A, Sztolsztener ME, Chacinska A** (2013) The ubiquitin-proteasome system regulates mitochondrial intermembrane space proteins. *Mol Cell Biol* **33**: 2136–2148
- Bragoszewski P, Wasilewski M, Sakowska P, Gornicka A, Böttinger L, Qiu J, Wiedemann N, Chacinska A** (2015) Retro-translocation of mitochondrial intermembrane space proteins. *Proc Natl Acad Sci U S A* **112**: 7713–7718
- Bräutigam A, Weber APM** (2009) Proteomic analysis of the proplastid envelope membrane provides novel insights into small molecule and protein transport across proplastid membranes. *Mol Plant* **2**: 1247–1261
- Bredemeier R, Schlegel T, Ertel F, Vojta A, Borissenko L, Bohnsack MT, Groll M, Von Haeseler A, Schleiff E** (2007) Functional and phylogenetic properties of the pore-forming β -barrel transporters of the Omp85 family. *J Biol Chem* **282**: 1882–1890

- Breuers FKH, Bräutigam A, Geimer S, Welzel UY, Stefano G, Renna L, Brandizzi F, Weber APM** (2012) Dynamic remodeling of the plastid envelope membranes - A tool for chloroplast envelope in vivo localizations. *Front Plant Sci* **3**: 7
- Broad W, Ling Q, Jarvis P** (2016) New insights into roles of ubiquitin modification in regulating plastids and other endosymbiotic organelles. *Int Rev Cell Mol Biol*. doi: 10.1016/bs.ircmb.2016.02.007
- Brouwer M, Ertel F, Mirus O, Koenig N, Flinner D, Bublak EJ, Tripp A, Hahn P, Tripp J, Hahn A, et al** (2012) Structure and conservation of the periplasmic targeting factor Tic22 protein from plants and cyanobacteria. *J Biol Chem* **287**: 24164–24173
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL** (2009) BLAST plus: architecture and applications. *BMC Bioinformatics* **10**: 1
- Carvalho P, Goder V, Rapoport TA** (2006) Distinct ubiquitin-Ligase complexes define convergent pathways for the degradation of ER proteins. *Cell* **126**: 361–373
- Chen Y-L, Chen L-J, Li H** (2016) Polypeptide transport-associated domains of the Toc75 channel protein are located in the intermembrane space of chloroplasts. *Plant Physiol* **172**: 235–243
- Chou M-L, Fitzpatrick LM, Tu S-L, Budziszewski G, Potter-Lewis S, Akita M, Levin JZ, Keegstra K, Li H-M** (2003) Tic40, a membrane-anchored co-chaperone homolog in the chloroplast protein translocon. *EMBO J* **22**: 2970–2980
- Chu CC, Li HM** (2011) Determining the location of an arabidopsis chloroplast protein using in vitro import followed by fractionation and alkaline extraction. *Methods Mol Biol* **774**: 339–350
- Clark KA, Krysan PJ** (2010) Chromosomal translocations are a common phenomenon in *Arabidopsis thaliana* T-DNA insertion lines. *Plant J* **64**: 990–1001

- Constan D, Patel R, Keegstra K, Jarvis P** (2004) An outer envelope membrane component of the plastid protein import apparatus plays an essential role in Arabidopsis. *Plant J* **38**: 93–106
- Dahlin C, Cline K** (1991) Developmental regulation of the plastid protein import apparatus. *Plant Cell* **3**: 1131–1140
- Dastvan R, Brouwer E-M, Schuetz D, Mirus O, Schleiff E, Prisner TF** (2016) Relative orientation of POTRA domains from cyanobacterial Omp85 studied by pulsed EPR spectroscopy. *Biophys J* **110**: 2195–206
- Day PM, Potter D, Inoue K** (2014) Evolution and targeting of Omp85 homologs in the chloroplast outer envelope membrane. *Front Plant Sci* **5**: 535
- Delattre A-SS, Clantin B, Saint N, Locht C, Villeret V, Jacob-Dubuisson F** (2010) Functional importance of a conserved sequence motif in FhaC, a prototypic member of the TpsB/Omp85 superfamily. *FEBS J* **277**: 4755–4765
- Demarsy E, Lakshmanan AM, Kessler F** (2014) Border control: selectivity of chloroplast protein import and regulation at the TOC-complex. *Front Plant Sci* **5**: 483
- Deshaies RJ, Joazeiro CAP** (2009) RING domain E3 ubiquitin ligases. *Annu Rev Biochem* **78**: 399–434
- Dhanoo PK, Richardson LGL, Smith MD, Gidda SK, Henderson MPA, Andrews DW, Mullen RT** (2010) Distinct pathways mediate the sorting of tail-anchored proteins to the plastid outer envelope. *PLoS One* **5**: e10098
- Dubin MJ, Bowler C, Benvenuto G, Clough S, Bent A, Hartley J, Temple G, Brasch M, Karimi M, Inze D, et al** (2008) A modified Gateway cloning strategy for overexpressing tagged proteins in plants. *Plant Methods* **4**: 3
- Dutta S, Mohanty S, Tripathy BC** (2009) Role of temperature stress on chloroplast biogenesis and protein import in pea. *Plant Physiol* **150**: 1050–1061

- Dutta S, Teresinski HJ, Smith MD, Troesch R, Vidi P** (2014) A Split-Ubiquitin yeast two-hybrid screen to examine the substrate specificity of atToc159 and atToc132, two Arabidopsis chloroplast preprotein import receptors. *PLoS One* **9**: e95026
- Endow JK, Rocha AG, Baldwin AJ, Roston RL, Yamaguchi T, Kamikubo H, Inoue K** (2016) Polyglycine acts as a rejection signal for protein transport at the chloroplast envelope. *PLoS One* **11**: e0167802
- Ertel F, Mirus O, Bredemeier R, Moslavac S, Becker T, Schleiff E** (2005) The evolutionarily related β -barrel polypeptide transporters from *Pisum sativum* and *Nostoc PCC7120* contain two distinct functional domains. *J Biol Chem* **280**: 28281–28289
- Flores-Pérez Ú, Jarvis P** (2013) Molecular chaperone involvement in chloroplast protein import. *Biochim Biophys Acta - Mol Cell Res* **1833**: 332–340
- Froehlich J** (2011) Studying arabidopsis envelope protein localization and topology using thermolysin and trypsin proteases. *Methods Mol. Biol. Humana Press, Clifton, N.J.*, pp 351–367
- Golczyk H, Greiner S, Wanner G, Weihe A, Bock R, Borner T, Herrmann RG** (2014) Chloroplast DNA in mature and senescing leaves: A reappraisal. *Plant Cell* **26**: 847–854
- Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, et al** (2012) Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Res* **40**: D1178-D1186
- Gould SB, Maier U-G, Martin WF** (2015) Protein import and the origin of red complex plastids. *Curr Biol* **25**: 515–521
- Gu Y, Li H, Dong H, Zeng Y, Zhang Z, Paterson NG, Stansfeld PJ, Wang Z, Zhang Y, Wang W, et al** (2016) Structural basis of outer membrane protein insertion by the BAM complex. *Nature* **531**: 64–69

- Guerra DD, Callis J** (2012) Update on the ubiquitin modification system
Ubiquitin on the move: The ubiquitin modification system plays diverse roles in the regulation of endoplasmic reticulum and plasma membrane-localized proteins. *Plant Physiol* **160**: 56–64
- Gutensohn M, Schulz B, Nicolay P, Flügge UI** (2000) Functional analysis of the two Arabidopsis homologues of Toc34, a component of the chloroplast protein import apparatus. *Plant J* **23**: 771–783
- Heinz E, Lithgow T** (2014) A comprehensive analysis of the Omp85/TpsB protein superfamily structural diversity, taxonomic occurrence, and evolution. *Front Microbiol* **5**: 1–13
- Hiltbrunner A, Bauer J, Vidi PA, Infanger S, Weibel P, Hohwy M, Kessler F** (2001) Targeting of an abundant cytosolic form of the protein import receptor at Toc159 to the outer chloroplast membrane. *J Cell Biol* **154**: 309–316
- Holt MT, David Y, Pollock S, Tang Z, Jeon J, Kim J, Roeder RG, Muir TW** (2015) Identification of a functional hotspot on ubiquitin required for stimulation of methyltransferase activity on chromatin. *Proc Natl Acad Sci U S A* **112**: 10365–10370
- Hooper CM, Tanz SK, Castleden IR, Vacher MA, Small ID, Millar AH** (2014) SUBAcon: A consensus algorithm for unifying the subcellular localization data of the Arabidopsis proteome. *Bioinformatics* **30**: 3356–3364
- Hsu S-C, Patel R, Bédard J, Jarvis P, Inoue K** (2008) Two distinct Omp85 paralogs in the chloroplast outer envelope membrane are essential for embryogenesis in Arabidopsis thaliana. *Plant Signal Behav* **3**: 1134–1135
- Hsueh YC, Flinner N, Gross LE, Haarmann R, Mirus O, Sommer MS, Schleiff E** (2017) The chloroplast outer envelope protein P39 in Arabidopsis thaliana belongs to the Omp85 protein family. *Proteins* **85**: 1391–1401
- Hua Z, Vierstra RD** (2016) Ubiquitin goes green. *Trends Cell Biol* **26**: 3–5

- Huang W, Ling Q, Bédard J, Lilley K, Jarvis P** (2011) In vivo analyses of the roles of essential Omp85-related proteins in the chloroplast outer envelope membrane. *Plant Physiol* **157**: 147–159
- Infanger S, Bischof S, Hiltbrunner A, Agne B, Baginsky S, Kessler F** (2011) The chloroplast import receptor Toc90 partially restores the accumulation of Toc159 client proteins in the arabidopsis thaliana ppi2 mutant. *Mol Plant* **4**: 252–263
- Inoue K, Demel R, de Kruijff B, Keegstra K** (2001) The N-terminal portion of the preToc75 transit peptide interacts with membrane lipids and inhibits binding and import of precursor proteins into isolated chloroplasts. *Eur J Biochem* **268**: 4036–4043
- Inoue K, Keegstra K** (2003) A polyglycine stretch is necessary for proper targeting of the protein translocation channel precursor to the outer envelope membrane of chloroplasts. *Plant J* **34**: 661–669
- Ishida H, Izumi M, Wada S, Makino A** (2014) Roles of autophagy in chloroplast recycling. *Biochim Biophys Acta - Bioenerg* **1837**: 512–521
- Ivanova Y, Smith MD, Chen K, Schnell DJ** (2004) Members of the Toc159 import receptor family represent distinct pathways for protein targeting to plastids. *Mol Biol Cell* **15**: 3379–3392
- Izumi M, Hidema J, Ishida H** (2015) From Arabidopsis to cereal crops: Conservation of chloroplast protein degradation by autophagy indicates its fundamental role in plant productivity. *Plant Signal Behav* **10**: 10–12
- Jarvis P** (2008) Targeting of nucleus-encoded proteins to chloroplasts in plants. *New Phytol* **179**: 257–285
- Jarvis P, Chen L-J, Li H, Peto CA, Fankhauser C, Chory J** (1998) An Arabidopsis mutant defective in the plastid general protein import apparatus. *Science* **282**: 100–103
- Jarvis P, López-Juez E** (2013) Biogenesis and homeostasis of chloroplasts and other plastids. *Nat Rev Mol Cell Biol* **14**: 787–802

- Jelic M, Soll J, Schleiff E** (2003) Two Toc34 homologues with different properties. *Biochemistry* **42**: 5906–5916
- Jia F, Wang C, Huang J, Yang G, Wu C, Zheng C** (2015) SCF E3 ligase PP2-B11 plays a positive role in response to salt stress in Arabidopsis. *J Exp Bot* **66**: 4683–4697
- Jouhet J, Gray JC** (2009) Interaction of actin and the chloroplast protein import apparatus. *J Biol Chem* **284**: 19132–19141
- Kakizaki T, Matsumura H, Nakayama K, Che F-S, Terauchi R, Inaba T** (2009) Coordination of plastid protein import and nuclear gene expression by plastid-to-nucleus retrograde signaling. *Plant Physiol* **151**: 1339–1353
- Kalanon M, McFadden GI** (2008) The chloroplast protein translocation complexes of *Chlamydomonas reinhardtii*: A bioinformatic comparison of Toc and Tic components in plants, green algae and red algae. *Genetics* **179**: 95–112
- Kasmati AR, Töpel M, Khan NZ, Patel R, Ling Q, Karim S, Aronsson H, Jarvis P** (2013) Evolutionary, molecular and genetic analyses of Tic22 homologues in *Arabidopsis thaliana* chloroplasts. *PLoS One* **8**: e63863
- Katoh K, Standley DM** (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol* **30**: 772–780
- Kessler F** (2012) Chloroplast delivery by UPS: Identification of a membrane-anchored E3 ligase in plants reveals a role for the ubiquitin proteasome system in chloroplast development. *Science* **338**: 622–623
- Kikuchi S, Bédard J, Hirano M, Hirabayashi Y, Oishi M, Imai M, Takase M, Ide T, Nakai M** (2013) Uncovering the protein translocon at the chloroplast inner envelope membrane. *Science* **339**: 571–574
- Kikuchi S, Hirohashi T, Nakai M** (2006) Characterization of the preprotein translocon at the outer envelope membrane of chloroplasts by Blue Native PAGE. *Plant Cell Physiol* **47**: 363–371

- Kim D-Y, Scalf M, Smith LM, Vierstra RD** (2013) Advanced proteomic analyses yield a deep catalog of ubiquitylation targets in Arabidopsis. *Plant Cell* **25**: 1523–1540
- Kleiger G, Mayor T** (2014) Perilous journey: A tour of the ubiquitin-proteasome system. *Trends Cell Biol* **24**: 352–359
- Kovacheva S, Bédard J, Patel R, Dudley P, Twell D, Ríos G, Koncz C, Jarvis P** (2005) In vivo studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. *Plant J* **41**: 412–428
- Kovacheva S, Bedard J, Wardle A, Patel R, Jarvis P** (2007) Further in vivo studies on the role of the molecular chaperone , Hsp93 , in plastid protein import. *Plant J* **50**: 364–379
- Kubis S, Baldwin A, Patel R, Razzaq A, Dupree P, Lilley K, Kurth J, Leister D, 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, Bédard J, Kovacheva S, Lilley K, Biehl A, Leister D, Ríos G, Koncz C, et al** (2004) Functional specialization amongst the Arabidopsis Toc159 family of chloroplast protein import receptors. *Plant Cell* **16**: 2059–2077
- Kubis SE, Lilley KS, Jarvis P** (2008) Isolation and preparation of chloroplasts from Arabidopsis thaliana plants. *Methods Mol. Biol. Humana Press, Clifton, N.J.*, pp 171–186
- Laemmli UK** (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680–5
- Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DL, Garcia-Hernandez M, et al** (2012) The Arabidopsis Information Resource (TAIR): Improved gene annotation and new tools. *Nucleic Acids Res* **40**: D1202–D1210

- Lee J, Kim DH, Hwang I** (2014) Specific targeting of proteins to outer envelope membranes of endosymbiotic organelles, chloroplasts, and mitochondria. *Front Plant Sci* **5**: 173
- Lee S, Lee DW, Lee Y, Mayer U, Stierhof Y-D, Lee S, Jurgens G, Hwang I** (2009) Heat shock protein cognate 70-4 and an E3 ubiquitin ligase, CHIP, mediate plastid-destined precursor degradation through the ubiquitin-26S proteasome system in Arabidopsis. *Plant Cell* **21**: 3984–4001
- Lemus L, Goder V** (2014) Regulation of endoplasmic reticulum-associated protein degradation (ERAD) by ubiquitin. *Cells* **3**: 824–847
- Li F, Vierstra RD** (2012) Autophagy: a multifaceted intracellular system for bulk and selective recycling. *Trends Plant Sci* **17**: 526–537
- Li G, Froehlich JE, Elowsky C, Msanne J, Ostosh AC, Zhang C, Awada T, Alfano JR** (2014) Distinct *Pseudomonas* type-III effectors use a cleavable transit peptide to target chloroplasts. *Plant J* **77**: 310–321
- Li H, Chiu C-C** (2010) Protein Transport into Chloroplasts. *Annu Rev Plant Biol* **61**: 157–180
- Li H min, Teng YS** (2013) Transit peptide design and plastid import regulation. *Trends Plant Sci* **18**: 360–366
- Li M, Rong Y, Chuang Y-S, Peng D, Emr SD** (2015) Ubiquitin-dependent lysosomal membrane protein sorting and degradation. *Mol Cell* **57**: 467–478
- Ling Q, Huang W, Baldwin A, Jarvis P** (2012) Chloroplast biogenesis is regulated by direct action of the ubiquitin-proteasome system. *Science* **338**: 655–659
- Ling Q, Huang W, Jarvis P** (2011) Use of a SPAD-502 meter to measure leaf chlorophyll concentration in Arabidopsis thaliana. *Photosynth Res* **107**: 209–214
- Ling Q, Jarvis P** (2015a) Functions of plastid protein import and the ubiquitin-proteasome system in plastid development. *Biochim Biophys Acta* **1847**: 939–48

- Ling Q, Jarvis P** (2015b) Regulation of chloroplast protein import by the ubiquitin E3 ligase SP1 is important for stress tolerance in plants. *Curr Biol* **25**: 2527–2534
- Ling Q, Jarvis P** (2016) Plant signaling: ubiquitin pulls the trigger on chloroplast degradation. *Curr Biol* **26**: R38–R40
- López-Juez E, Hills A** (2011) Screening or selection for chloroplast biogenesis mutants of *Arabidopsis*, following chemical or insertional mutagenesis. *Methods Mol Biol* **774**: 3–18
- Maier T, Clantin B, Gruss F, Dewitte F, Delattre A-S, Jacob-Dubuisson F, Hiller S, Villeret V** (2015) Conserved Omp85 lid-lock structure and substrate recognition in FhaC. *Nat Commun* **6**: 7452
- Matsushima R, Hu Y, Toyoda K, Sodmergen, Sakamoto W** (2008) The model plant *Medicago truncatula* exhibits biparental plastid inheritance. *Plant Cell Physiol* **49**: 81–91
- May T, Soll J** (1998) Positive charges determine the topology and functionality of the transmembrane domain in the chloroplastic outer envelope protein Toc34. *J Cell Biol* **141**: 895–904
- McCuaig B, Dufour SC, Raguso RA, Bhatt AP, Marino P** (2015) Structural changes in plastids of developing *Splachnum ampullaceum* sporophytes and relationship to odour production. *Plant Biol* **17**: 466–473
- Mehnert M, Sommermeyer F, Berger M, Kumar Lakshmipathy S, Gauss R, Aebi M, Jarosch E, Sommer T** (2015) The interplay of Hrd3 and the molecular chaperone system ensures efficient degradation of malformed secretory proteins. *Mol Biol Cell* **26**: 185–194
- von Mikecz A, Chen M, Rockel T, Scharf A** (2008) The nuclear ubiquitin-proteasome system: visualization of proteasomes, protein aggregates, and proteolysis in the cell nucleus. *Methods Mol Biol* **463**: 191–202
- Misra R** (2012) Assembly of the β -Barrel Outer Membrane Proteins in Gram-Negative Bacteria, Mitochondria, and Chloroplasts. *ISRN Mol Biol* **2012**: 1–15

- Moslavac S, Mirus O, Bredemeier R, Soll J, Von Haeseler A, Schleiff E** (2005) Conserved pore-forming regions in polypeptide-transporting proteins. *FEBS J* **272**: 1367–1378
- Nakai M** (2015) 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
- Nakamura M, Toyota M, Tasaka M, Morita MT** (2011) An Arabidopsis E3 ligase, SHOOT GRAVITROPISM9, modulates the interaction between statoliths and F-actin in gravity sensing. *Plant Cell* **23**: 1830–1848
- Neff MM, Neff JD, Chory J, Pepper AE** (1998) dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: Experimental applications in Arabidopsis thaliana genetics. *Plant J* **14**: 387–392
- Nicolaisen K, Missbach S, Hsueh Y-C, Ertel F, Fulgosi H, Sommer MS, Schleiff E** (2015) The Omp85-type outer membrane protein P36 of Arabidopsis thaliana evolved by recent gene duplication. *J Plant Res* **128**: 317–325
- Noinaj N, Kuszak AJ, Balusek C, Gumbart JC, Buchanan SK** (2014) Lateral opening and exit pore formation are required for BamA function. *Structure* **22**: 1055–1062
- Noinaj N, Kuszak AJ, Gumbart JC, Lukacik P, Chang H, Easley NC, Lithgow T, Buchanan SK** (2013) Structural insight into the biogenesis of β -barrel membrane proteins. *Nature* **501**: 385–390
- Noinaj N, Rollauer SE, Buchanan SK** (2015) The β -barrel membrane protein insertase machinery from Gram-negative bacteria. *Curr Opin Struct Biol* **31**: 35–42
- O'Neil PK, Richardson LGL, Paila YD, Piszczek G, Chakravarthy S, Noinaj N, Schnell D** (2017) The POTRA domains of Toc75 exhibit chaperone-like function to facilitate import into chloroplasts. *Proc Natl Acad Sci U S A* **114**: E4868–E4876

- Oh YJ, Hwang I, Jun Oh Y, Hwang I** (2015) Targeting and biogenesis of transporters and channels in chloroplast envelope membranes: Unsolved questions. *Cell Calcium* **58**: 122–130
- Paila YD, Richardson LGL, Inoue H, Parks ES, McMahon J, Inoue K, Schnell DJ** (2016) Multi-functional roles for the polypeptide transport associated domains of Toc75 in chloroplast protein import. *eLife* **5**: 235–245
- Paila YD, Richardson LGL, Schnell DJ** (2015) New insights into the mechanism of chloroplast protein import and its integration with protein quality control, organelle biogenesis and development. *J Mol Biol* **427**: 1038–1060
- Patel R, Hsu S-C, Bedard J, Inoue K, Jarvis P** (2008) The Omp85-related chloroplast outer envelope protein OEP80 is essential for viability in *Arabidopsis*. *Plant Physiol* **148**: 235–245
- Petre B, Saunders DG, Sklenar J, Lorrain C, Win J, Duplessis S, Kamoun S** (2015) Candidate effector proteins of the rust pathogen *Melampsora larici-populina* target diverse plant cell compartments. *Mol Plant Microbe Interact* **28**: 689–700
- Plösch M, Reisinger V, Eichacker LA** (2011) Proteomic comparison of etioplast and chloroplast protein complexes. *J Proteomics* **74**: 1256–1265
- Pogson BJ, Ganguly D, Albrecht-Borth V** (2015) Insights into chloroplast biogenesis and development. *Biochim Biophys Acta* **1847**: 1017–1024
- Porra RJ, Thompson WA, Kriedemann PE** (1989) Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochim Biophys Acta - Bioenerg* **975**: 384–394
- Quilichini TD, Douglas CJ, Samuels AL** (2014) New views of tapetum ultrastructure and pollen exine development in *Arabidopsis thaliana*. *Ann Bot* **114**: 1189–1201

- Reynolds PHS, Blevins DG** (1986) Phosphoserine aminotransferase in Soybean root-nodules - Demonstration and localization. *Plant Physiol* **81**: 293–296
- Richardson LGL, Paila YD, Siman SR, Chen Y, Smith MD, Schnell DJ** (2014) Targeting and assembly of components of the TOC protein import complex at the chloroplast outer envelope membrane. *Front Plant Sci* **5**: 269
- Richly H, Rape M, Braun S, Rumpf S, Hoege C, Jentsch S** (2005) A series of ubiquitin binding factors connects CDC48/p97 to substrate multiubiquitylation and proteasomal targeting. *Cell* **120**: 73–84
- Richter S, Lamppa GK** (2003) Structural properties of the chloroplast stromal processing peptidase required for its function in transit peptide removal. *J Biol Chem* **278**: 39497–39502
- Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP** (2012) Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* **61**: 539–542
- Rudolf M, MacHettira AB, Groß LE, Weber KL, Bolte K, Bionda T, Sommer MS, Maier UG, Weber APM, Schleiff E, et al** (2013) In vivo function of Tic22, a protein import component of the intermembrane space of chloroplasts. *Mol Plant* **6**: 817–829
- Ruggiano A, Foresti O, Carvalho P** (2014) ER-associated degradation: Protein quality control and beyond. *J Cell Biol* **204**: 869–879
- Sadanandom A, Bailey M, Ewan R, Lee J, Nelis S** (2012) The ubiquitin-proteasome system: Central modifier of plant signalling. *New Phytol* **196**: 13–28
- Sakamoto W, Miyagishima S-Y, Jarvis P** (2008) Chloroplast biogenesis: control of plastid development, protein import, division and inheritance. *The Arabidopsis Book* **6**: e0110

- Sako K, Yanagawa Y, Kanai T, Sato T, Seki M, Fujiwara M, Fukao Y, Yamaguchi J** (2014) Proteomic analysis of the 26S proteasome reveals its direct interaction with transit peptides of plastid protein precursors for their degradation. *J Proteome Res* **13**: 3223–3230
- Sato EM, Hijazi H, Bennett MJ, Vissenberg K, Swarup R** (2015) New insights into root gravitropic signalling. *J Exp Bot* **66**: 2155–2165
- Schleiff E, Soll J, K  chler M, K  hlbrandt W, Harrer R** (2003) Characterization of the translocon of the outer envelope of chloroplasts. *J Cell Biol* **160**: 541–551
- Schwuchow JM, Kern VD, White NJ, Sack FD** (2002) Conservation of the plastid sedimentation zone in all moss genera with known gravitropic protonemata. *J Plant Growth Regul* **21**: 146–155
- Shao N, Krieger-Liszkay A, Schroda M, Beck CF** (2007) A reporter system for the individual detection of hydrogen peroxide and singlet oxygen: its use for the assay of reactive oxygen species produced in vivo. *Plant J* **50**: 475–487
- Shapiguzov A, Vainonen JP, Wrzaczek M, Kangasj  rvi J** (2012) ROS-talk – how the apoplast, the chloroplast, and the nucleus get the message through. *Front Plant Sci* **3**: 292
- Sharma P, Jha AB, Dubey RS, Pessarakli M** (2012) Reactive oxygen species, oxidative damage, and antioxidative defense mechanism in plants under stressful conditions. *J Bot* **2012**: 1–26
- Shen G, Adam Z, Zhang H** (2007) The E3 ligase AtCHIP ubiquitylates FtsH1, a component of the chloroplast FtsH protease, and affects protein degradation in chloroplasts. *Plant J* **52**: 309–321
- Shi L-X, Theg SM** (2013a) Energetic cost of protein import across the envelope membranes of chloroplasts. *Proc Natl Acad Sci U S A* **110**: 930–935
- Shi LX, Theg SM** (2013b) The chloroplast protein import system: From algae to trees. *Biochim Biophys Acta - Mol Cell Res* **1833**: 314–331

- Simm S, Papasotiriou DG, Ibrahim M, Leisegang MS, Müller B, Schorge T, Karas M, Mirus O, Sommer MS, Schleiff E** (2013) Defining the core proteome of the chloroplast envelope membranes. *Front Plant Sci* **4**: 11
- Sinnige T, Weingarth M, Renault M, Baker L, Tommassen J, Baldus M** (2014) Solid-State NMR studies of full-length BamA in lipid bilayers suggest limited overall POTRA mobility. *J Mol Biol* **426**: 2009–2021
- Sommer M, Rudolf M, Tillmann B, Tripp J, Sommer MS, Schleiff E** (2013) Toc33 and Toc64-III cooperate in precursor protein import into the chloroplasts of *Arabidopsis thaliana*. *Plant Cell Environ* **36**: 970–983
- Stamatakis A** (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313
- Stanga JP, Boonsirichai K, Sedbrook JC, Otegui MS, Masson PH** (2009) A role for the TOC complex in *Arabidopsis* root gravitropism. *Plant Physiol* **149**: 1896–1905
- Stegmeier JF, Andersen C** (2006) Characterization of pores formed by YaeT (Omp85) from *Escherichia coli*. *J Biochem* **140**: 275–283
- Stolz A, Schweizer RS, Schäfer A, Wolf DH** (2010) Dfm1 forms distinct complexes with Cdc48 and the ER ubiquitin ligases and is required for ERAD. *Traffic* **11**: 1363–1369
- Stroud DA, Becker T, Qiu J, Stojanovski D, Pfannschmidt S, Wirth C, Hunte C, Guiard B, Meisinger C, Pfanner N, et al** (2011) Biogenesis of mitochondrial β -barrel proteins: the POTRA domain is involved in precursor release from the SAM complex. *Mol Biol Cell* **22**: 2823–2833
- Su P-H, Lai Y-H** (2017) A reliable and non-destructive method for monitoring the stromal pH in isolated chloroplasts using a fluorescent pH probe. *Front Plant Sci* **8**: 2079
- Suzuki N, Koussevitzky S, Mittler R, Miller G** (2012) ROS and redox signalling in the response of plants to abiotic stress. *Plant Cell Environ* **35**: 259–270

- Svozil J, Hirsch-Hoffmann M, Dudler R, Gruissem W, Baerenfaller K** (2014) Protein abundance changes and ubiquitylation targets identified after inhibition of the proteasome with Syringolin A. *Mol Cell Proteomics* **13**: 1523–1536
- Taylor EB, Rutter J** (2011) Mitochondrial quality control by the ubiquitin–proteasome system. *Biochem Soc Trans* **39**: 1509–1513
- Teng YS, Chan PT, Li H min** (2012) Differential age-dependent import regulation by signal peptides. *PLoS Biol* **10**: e1001416
- Thrower JS, Hoffman L, Rechsteiner M, Pickart CM, Amerik A, Swaminathan S, Krantz B, Wilkinson K, Hochstrasser M, Appleman J, et al** (2000) Recognition of the polyubiquitin proteolytic signal. *EMBO J* **19**: 94–102
- Töpel M, Ling Q, Jarvis P** (2012) Neofunctionalization within the Omp85 protein superfamily during chloroplast evolution. *Plant Signal Behav* **7**: 161–164
- de Torres Zabala M, Littlejohn G, Jayaraman S, Studholme D, Bailey T, Lawson T, Tillich M, Licht D, Bölter B, Delfino L, et al** (2015) Chloroplasts play a central role in plant defence and are targeted by pathogen effectors. *Nat Plants* **1**: 15074
- Tsai J-YY, Chu C-CC, Yeh Y-HH, Chen L-JJ, Li HM, Hsiao C-DD** (2013) Structural characterizations of the chloroplast translocon protein Tic110. **75**: 847–857
- Tu S-L, Chen L-J, Smith MD, Su Y-S, Schnell DJ, Li H-M** (2004) Import pathways of chloroplast interior proteins and the outer-membrane protein OEP14 converge at Toc75. *Plant Cell* **16**: 2078–2088
- Ulrich T, Gross LE, Sommer MS, Schleiff E, Rapaport D** (2012) Chloroplast β -barrel proteins are assembled into the mitochondrial outer membrane in a process that depends on the TOM and TOB complexes. *J Biol Chem* **287**: 27467–27479

- Veen M, Stahl U, Lang C** (2003) Combined overexpression of genes of the ergosterol biosynthetic pathway leads to accumulation of sterols in. *FEMS Yeast Res* **4**: 87–95
- Vembar SS, Brodsky JL** (2008) One step at a time: endoplasmic reticulum-associated degradation. *Nat Rev Mol Cell Biol* **9**: 944–957
- Vierstra RD** (2009) The ubiquitin–26S proteasome system at the nexus of plant biology. *Nat Rev Mol Cell Biol* **10**: 385–397
- Villarejo A, Burén S, Larsson S, Déjardin A, Monné M, Rudhe C, Karlsson J, Jansson S, Lerouge P, Rolland N, et al** (2005) Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. *Nat Cell Biol* **7**: 1224–1231
- Wada S, Ishida H, Izumi M, Yoshimoto K, Ohsumi Y, Mae T, Makino A** (2009) Autophagy plays a role in chloroplast degradation during senescence in individually darkened leaves. *Plant Physiol* **149**: 885–893
- Wallas TR, Smith MD, Sanchez-Nieto S, Schnell DJ** (2003) The roles of Toc34 and Toc75 in targeting the Toc159 preprotein receptor to chloroplasts. *J Biol Chem* **278**: 44289–44297
- Walsh CK, Sadanandom A** (2014) Ubiquitin chain topology in plant cell signaling: a new facet to an evergreen story. *Plant Genet Genomics* **5**: 122
- Wan J, Blakeley SD, Dennis DT, Ko K** (1996) Transit peptides play a major role in the preferential import of proteins into leucoplasts and chloroplasts. *J Biol Chem* **271**: 31227–31233
- Wang F, Agne B, Kessler F, Schnell DJ** (2008) The role of GTP binding and hydrolysis at the atToc159 preprotein receptor during protein import into chloroplasts. *J Cell Biol* **183**: 87–99
- Weigel D, Glazebrook J** (2006) Transformation of *Agrobacterium* using the freeze-thaw method. *Cold Spring Harb Protoc* **2006**: pdb.prot4666
- Wise RR** (2006) The diversity of plastid form and function. *In* R Wise, K Hooper, eds, *Struct. Funct. Plast.* Springer Netherlands, Dordrecht, pp 3–26

- Wolf DH, Stolz A** (2012) The Cdc48 machine in endoplasmic reticulum associated protein degradation. *Biochim Biophys Acta - Mol Cell Res* **1823**: 117–124
- Woodson JD, Joens MS, Sinson AB, Gilkerson J, Salomé PA, Weigel D, Fitzpatrick JA, Chory J** (2015) Ubiquitin facilitates a quality-control pathway that removes damaged chloroplasts. *Science* **350**: 450–454
- Wu F-H, Shen S-C, Lee L-Y, Lee S-H, Chan M-T, Lin C-S** (2009) Tape-Arabidopsis Sandwich - a simpler Arabidopsis protoplast isolation method. *Plant Methods* **5**: 16
- Wu X, Li L, Jiang H** (2016) Doa1 targets ubiquitinated substrates for mitochondria-associated degradation. *J Cell Biol* **213**: 49–63
- Xie Q, Michaeli S, Peled-Zehavi H, Galili G** (2015) Chloroplast degradation: One organelle, multiple degradation pathways. *Trends Plant Sci* **20**: 264–265
- Yan X, Khan S, Hase T, Emes MJ, Bowsher CG** (2006) Differential uptake of photosynthetic and non-photosynthetic proteins by pea root plastids. *FEBS Lett* **580**: 6509–6512
- Yang Q, Chen ZZ, Zhou XF, Yin HB, Li X, Xin XF, Hong XH, Zhu JK, Gong Z** (2009) Overexpression of SOS (salt overly sensitive) genes increases salt tolerance in transgenic Arabidopsis. *Mol Plant* **2**: 22–31
- Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL** (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* **13**: 134
- Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D** (2004) A Molecular timeline for the origin of photosynthetic eukaryotes. *Mol Biol Evol* **21**: 809–818
- Yu T-S, Li H** (2001) Chloroplast protein translocon components atToc159 and atToc33 are not essential for chloroplast biogenesis in guard cells and root cells. *Plant Physiol* **127**: 90–96

- Zhang T, Ye Y** (2016) Doal is a MAD adaptor for Cdc48. *J Cell Biol* **213**: 7–9
- Zimorski V, Ku C, Martin WF, Gould SB** (2014) Endosymbiotic theory for organelle origins. *Curr Opin Microbiol* **22**: 38–48
- Zimorski V, Martin WF** (2014) Subcellular targeting of proteins and pathways during evolution. *New Phytol* **201**: 1–2
- Zufferey M, Montandon C, Douet V, Demarsy E, Agne B, Baginsky S, Kessler F** (2017) The novel chloroplast outer membrane kinase KOC1 is a required component of the plastid protein import machinery. *J Biol Chem* **292**: 6952–6964