UNIVERSITÉ DE NEUCHÂTEL

Discovering The Functions of The Acidic Domain of Toc159

A dissertation submitted to the

UNIVERSITY OF NEUCHATEL

for the degree of

Doctor of Natural Sciences

Submitted by

Ashok MUNUSAMY LAKSHMANAN

Accepted on the recommendation of

Prof. Felix Kessler, Thesis director

Prof. Jean-Marc Neuhaus

Prof. Samuel Zeeman

Prof. Paul Jarvis

2014

UNIVERSITÉ DE NEUCHÂTEL

Faculté des sciences
Secrétariat-décanat de Faculté
Rue Emile-Argand 11
2000 Neuchâtel - Suisse
Tél: + 41 (0)32 718 2100
E-mail: secretariat.sciences@unine.ch

IMPRIMATUR POUR THESE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise l'impression de la présente thèse soutenue par

Monsieur Ashok MUNUSAMY LAKSHMANAN

Titre:

"Discovering the Functions of the Acidic Domain of Toc159"

sur le rapport des membres du jury composé comme suit:

- Prof. Felix Kessler, Université de Neuchâtel, directeur de thèse
- Prof. Jean-Marc Neuhaus, Université de Neuchâtel
- Prof. Samuel C. Zeeman, ETH, Zürich
- Prof. R. Paul Jarvis, University of Oxford, UK

Neuchâtel, le 16 octobre 2014 Le Doyen, Prof. B. Colbois

Index

Table of Figures

Table of Appendices

Abbreviations

αCBP	antibodies recognizing calmodulin binding protein in the TAP tag
αHis	antibodies recognizing His_6
αToc159A	antibodies recognizing Toc159A
αToc33	antibodies recognizing Toc33
A-domain	acidic domain of Toc159 (if not purposefuly generalised for Toc159 family)
APS	ammonium persulfate
BSA	bovine serum albumin
BN-PAGE	blue native polyacrylamide gel electrophoresis
DEAE	diethylaminoethyl-resin
DHFR	dihydrofolatereductase
DSP	dithiobis(succinimidylpropionate)
EDTA	ethylenediamine-N,N,N',N'-tetraacetic acid
HEPES	4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid
IgG	immunoglobuline G
IPTG	isopropyl-β-D-1-thiogalactopyranoside potassium acetate
Ni-NTA	nickel-nitrilotriacetic acid
PBS	salined phosphate buffer
PCR	polymerase chain reaction polyethyleneimine-cellulose
pE1α	precursor of pyruvate reductase subunit E1 alpha
ppi	plastid protein import
ppi1	*toc33* knockout mutant
ppi2	*toc159* knockout mutant
ppi3	*toc132* knockout mutant
pRBCS	precursor of the small subunit of RubisCO
pSSU	precursor of the small subunit of RubisCO
pTic40	precursor of Tic40
RUBisCO	ribulose-1,5-bisphosphate carboxylase oxygenase
SDS	sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
TAP	Tandem Affinity Purification
TBS	salined Tris buffer
TEMED	N,N,N',N',-tetramethyl-ethylenediamine
Tic	translocon at the inner chloroplast membrane transmembrane helix
Toc	translocon at the outer chloroplast membrane
Tris	tris(hydroxymethyl)aminomethane
v/v	volume per volume
wt	wildtype
w/v	weight per volume
w/o ST	without amino acids serine and theronine
YFP	yellow fluorescence protein

Upper case, italic	gene, wild type allele (e.g. *TOC159*)
Lower case, italic	mutant allle (e.g. *ppi2*)
Regular case	protein (e.g. Toc159A)

atToc.., atTic	indicates the Toc and Tic protein of *A.thaliana*

Abstract

Preprotein import from the cytosol is a key mechanism in chloroplast biogenesis. Hundreds of different preproteins take an import route involving the chloroplast surface receptor Toc159. Toc159 has three domains A (acidic), G (GTPase) and M (membrane insertion). The N-terminal, cytosolic A-domain has been shown to contribute to preprotein specificity of Toc159 (Toc159A). This thesis began with specific questions to address the existence of hyperphosphorylated and soluble Toc159A as a separate protein in the cytosol and aims at expanding our understanding of Toc159-dependent chloroplast protein import

In its first part, this thesis demonstrates the dual localization of Toc159A both in the cytosol and at the chloroplast envelope using TAP (tandem affinity purification)-tagged Toc159A and Toc159A-YFP (yellow fluorescent protein). Blue native polyacrylamide gel electrophoresis BN-PAGE in combination with Western blotting identified Toc159A-TAP complexes. However, the weak stability of the complexes required stabilization of the complexes by chemical crosslinking by formaldehyde prior to purification. Complexes identified by BN-PAGE were confirmed by formaldeyhde crosslinking. Complexes were affinity purified and analysed by mass spectrometry. We identified putative interacting proteins that could be classified into three main categories. 1) Chloroplast targeted Toc159-dependent preproteins 2) two kinases predicted to be located at the chloroplast envelope and in the cytoplasm, 3) cytosolically localized Hsp70 and Calmodulin. Upon confirmation, all of these offer exciting future research perspectives.

The second part of the thesis focuses on *in vitro* experiments to understand possible roles of Toc159A in chloroplast protein import and Toc159A phosphorylation. Two hypotheses were tested:

1) First hypothesis: direct binding of the pSSU (pre-small subunit of Rubisco) preprotein to Toc159A prior to import into the organelle. This hypothesis was tested by adding recombinant A-domain to in vitro import assays. Inhibition of import at the earliest stages suggested an effect on the kinetics of import.

2) Second hypothesis: regulation of ribosomal translational activity by the A-domain. This hypothesis was tested by adding phosphorylatable or non-phosphorylatable A-domain to in vitro translation reactions. Phosphorylatable Toc159A had a positive effect on the *in vitro* translation of an early developmental stage specific preprotein pSSU but negatively affected the late developmental stage specific preprotein pTic40. This suggests that Toc159 may directly affect the synthesis of preproteins that take an import route required early in development. While these results are in need of further experimentation, they suggest a novel mechanism for coupling preprotein translation to chloroplast protein import.

1. Introduction

Eukaryotic cells are composed of multiple compartments that acquire specialized sets of proteins for function. The vast majority of proteins are encoded by the nuclear genome. After synthesis in the cytosol accurate protein sorting and export towards their destination organelles rely on intrinsic topogenic sequences(Blobel, 1980). Initially, correct recognition of a preprotein requires specific receptors at the surface of the organelle. This crucial step of intracellular trafficking control can be viewed as a key-lock type mechanism.

Plant chloroplasts import impressive quantities as well as an enormous diversity of proteins from the cytosol. Large scale proteome studies indicate that 2000 to 4000 different proteins follow the chloroplast route (Ferro et al., 2003; Friso et al., 2004; Kleffmann et al., 2004; Leister, 2003a). In the cytosol, chloroplast proteins are generally synthesized as preproteins with a N terminal targeting sequence that is cleaved to produce the mature chloroplast protein upon import. This N-terminal targeting sequence, named transit peptide in the context of chloroplast protein import, faithfully guides the preprotein to the chloroplast surface where it engages the import machinery. In the following, the preprotein is translocated across the dual envelope membranes into the stroma. The transit sequence is cleaved upon arrival in the stroma yielding the mature form of the protein followed by folding in the stroma, targeting to the inner membrane via the conservative sorting pathway, or transport to the thylakoid membrane system. The recognition and translocation of the preprotein at the plastid envelope is provided by the TOC-TIC (Translocon of Outer membrane Complex, Translocon of Inner membrane Complex) import machinery. In pea, the core TOC complex consists of an assembly of the two GTP dependent receptors Toc34 and Toc159 together with the b-barrel protein conducting channel Toc75 (Becker et al., 2004; Hirsch et al., 1994; Kessler et al., 1994; Perry & Keegstra, 1994; Schnell et al., 1994). Upon engagement of the preprotein, the TOC complex associates with the TIC complex to form a continuous channel through the plastid envelope. The protein conducting channel at the TIC complex has been suggested to be made up of Tic110 or Tic20, or yet a combination of the two. Recently, however, it has been suggested that four core components form a 1MDa TIC channel (Tic20, Tic214 formerly known as YCF1, Tic56 and Tic100; (Kikuchi et al., 2013). Protein synthesis and targeting involve a large variety of cellular activities that are energy-requiring. Solely translocation of a single preprotein across the chloroplast envelope through the TOC-TIC machinery requires the hydrolysis of 650 ATP molecules on average, representing about 0.6% of the total light-saturated energy output of the organelles (Shi & Theg, 2013). Therefore a tight control of TOC-TIC mediated import activity is required to respect the cellular energy budget allocated to protein import.

Plants originate from a primary endosymbiotic event involving a photosynthetic cyanobacterium captured by a eukaryotic cell. The evolution of plants towards complex and multicellular organisms has been accompanied by the diversification of interconvertible plastid types displaying distinct and highly specialized biochemical and physiological functions (Jarvis & Lopez-Juez, 2013). For instance the most prominent plastid type, the chloroplast, develop from proplastid or partially differentiated, non-photosynthetic etioplast, and can also differentiate into other non-photosynthetic plastid types such as chromoplast or elaioplast. Each plastid type requires the import of different subsets of proteins (Kleffmann et al., 2007; Brautigam & Weber, 2009; Barsan et al., 2012). Several strategies have evolved coordinately to ensure the selective import of plastid proteins. Together with the defined regulation of preprotein availability at the transcriptional levels, evolution also triggered diversification and increased complexity of both preprotein transit sequences (Heijne & Nishikawa, 1991; Li & Teng, 2013b) and composition of the import machinery (Gross & Bhattacharya, 2009; Kalanon & McFadden, 2008; Reumann et al., 2005; Shi et al., 2013). Evidence for the existence of different isoforms of TOC complex components has now been reported for several higher plant species including Arabidopsis, pea and tomato (Chang et al., 2014; Jackson-Constan & Keegstra, 2001; Yan et al., 2014). Each isoform is thought to preferentially import a specific subset of client preproteins that may be the result of differential binding affinity (Dutta et al., 2014; Inoue et al., 2010; Jelic et al., 2003; Smith, 2004). Therefore, the relative abundance of Toc isoforms may reflect the protein composition of a given plastid type and be a key marker of plastid identity (Ling et al., 2012).

On top of that, plants are sessile organisms and need to adapt to ever-changing environmental conditions. Dynamic regulation of TOC complex composition may occur at the posttranslational level and represent a key regulatory mechanism contributing to the change in protein composition. By consequence this allows rapid modulation of plastid metabolism to ensure and drive plant development and acclimation. Thus, the relative abundance of Toc receptor may not only be a marker of plastid type but also of plastid state (Agne & Kessler, 2010; Ling et al., 2012).

The molecular mechanisms underlying the process of protein translocation have been reviewed extensively (Andrès et al., 2010; Jarvis, 2008; Li & Chiu, 2010). Here, we present the current knowledge with regard to the selectivity and the regulation of the preprotein import process at the level of the TOC complex.

1.1. Preprotein import in plastid is regulated by developmental and environmental factors

Years before the identification of any of the components of the chloroplast protein import machinery Dahlin and Cline proposed that import activity is correlated with protein demands during plastid development (Dahlin & Cline, 1991). They observed a high import activity in non-photosynthetic proplastids, which gradually decreased as plastids matured. This phenomenon was observed for etioplast as well as chloroplast development. Interestingly, when dark-grown plants were shifted from dark to light the import activity of etioplasts was activated to accommodate the set of preproteins required for chloroplast differentiation (Dahlin & Cline, 1991). This seminal study focused on a few substrates and, given the experimental limitations at the time, was unable to provide a complete picture of plastid protein import dynamics. Recently, this topic was reinvestigated using a larger number of chloroplasts precursors proteins (Teng et al., 2012). This study confirmed that preprotein specificity is modulated in synchrony with chloroplast developmental stages. Interestingly, this study demonstrated that the earlier results by Dahlin and Cline cannot be extended to all import substrates. Rather, Teng and colleagues refined the model and classified the substrates according to their importability in chloroplasts at different developmental stages and consequently defined three age-selective classes: substrates that are imported more efficiently in young chloroplasts (group I), in older chloroplasts (group III), whereas group II represents substrates that are imported similarly in developing and mature chloroplasts. Thus, it appears that regulation of chloroplast preprotein import is part of a differential age-specific regulatory network.

In vitro import experiments using different isolated plastid types as well as the visualization of protein targeting using transgenic lines expressing transit peptides fused to GFP support the notion that import selectivity is regulated in a tissue specific manner (Primavesi et al., 2008; Wan et al., 1996; Yan et al., 2014). Finally temperature stress (cold and heat) on intact pea leaves and isolated chloroplasts was found to reduce import of the small subunit of RubisCO preprotein (pSSU) (Dutta et al., 2009).

In summary, these results demonstrate that both plastid import activity and selectivity are modulated in accordance with plastid type, developmental stage, and environmental condition. For this purpose plants have evolved a complex set of preprotein import components with specialized features and regulatory mechanisms (Jarvis et al., 1998; Kubis et al., 2004).

1.2. Preprotein selectivity at the chloroplast import machinery

1.2.1. Overview of the TOC-TIC machinery

Translocon of outer membrane complex-translocon of inner membrane complex is the major protein import pathway in higher plants (Asano et al., 2004; Bauer et al., 2000; Kovacheva etal., 2005). Most of the proteins with cleavable transit peptides that are targeted to the stroma, thylakoid membranes and lumen follow this route, that is therefore vital for plastid biogenesis (Bischof et al., 2011; Dutta et al., 2014; Kessler & Schnell, 2006). The native TOC-TIC complex in pea and Arabidopsis has been found to include two GTPase-receptors Toc159 and Toc34, a channel protein Toc75 and at least three additional regulatory Toc proteins Toc64, Toc22 and Toc12 (Andrès et al., 2010). At the inner membrane at least 11 different proteins have been reported to be involved in the import process (Kikuchi et al., 2013; Kovacs-Bogdan et al., 2010). Electrophysiological experiments suggested that Tic110 and Tic20 could function as channels facilitating the translocation of preproteins across the inner membrane (Kikuchi et al., 2013; 2009; Kovács-Bogdán, et al., 2011). These two channels are thought to function independently and in different complexes (Kikuchi et al., 2009; 2013; Kovács-Bogdán et al., 2011). This is supported by the finding that Tic110 interacts with preproteins and TOC complexes (Inaba, 2005; Schnell et al., 1994) but not with Tic20 (Kikuchi et al., 2009). Tic110 is a protein of eukaryotic origin present in various plastid-containing organisms (Shi & Theg, 2013). Its function is indispensable for plant viability and chloroplast biogenesis (Inaba, 2005). Based on these data it was proposed that Tic110 has an essential role in chloroplast protein import. Recently, composition of the Tic20 complex in Arabidopsis has been investigated using Blue Native PAGE and mass spectrometric analyses. The results suggested that Tic20 associates with Tic56, Tic100 and Tic214 (Kikuchi et al., 2013). Although Tic20 is of prokaryotic origin and is well conserved among the plant kingdom, Tic56, Tic100 and Tic214 appear to have specifically evolved in a limited number of higher plant species only (Kikuchi et al., 2013). Tic214, also known as YCF1, is absent from the genome of some Poacae species (Jensen & Leister, 2014; Smith & Lee, 2014), thus the role of TIC20 complex as the general inner chloroplast membrane translocon in higher plants is questionable. Nevertheless the albino, seedling lethal phenotype of null mutants of each of the TIC20 complex subunits underscores their functional importance at least in Arabidopsis. In conclusion the exact contribution of TIC110 and TIC20 complexes in chloroplast protein import is still under debate.

At the evolutionary level, a view of growing complexity of the composition of TOC machinery is emerging (Shi & Theg, 2013b). Starting with one channel protein at the outer envelope in cyanobacteria, the outer envelope protein import complex has evolved into a GTP-regulated multi-protein complex in higher plants (Hiltbrunner et al., 2001; Kessler & Schnell, 2002; Olsen & Keegstra, 1992; Schnell et al., 1994; Voulhoux et

al., 2003). The Toc receptors can form homo- and heterodimers in a dynamic way regulated by preprotein binding and GTP binding/hydrolysis activity (Lee et al., 2009b; Oreb et al., 2011; Rahim et al, 2009; Smith, 2002; Sun et al., 2002; Wallas et al., 2003). Although GTP binding and GTPase activity seem dispensable (expression of GTPase/dimerization-defective Toc159 and Toc33 complement the corresponding knock out mutants), it has been shown that they are required for full preprotein import efficiency *in vitro* (Agne et al., 2009; Aronsson & Jarvis, 2011; Aronsson et al., 2010). In most higher plants the Toc75 channel is encoded by a single gene (Inoue & Keegstra, 2003), but normally more than one homolog for the plastid specific GTPase families Toc159 and Toc34 exists, and thus there is the possibility of making various combinations of TOC complexes (Chang et al., 2014; Hiltbrunner et al., 2001; Yan et al., 2014). The evolution of a translocation route depending on GTP-binding as well as other accessory proteins may be seen as the key to the developmental stage specific regulation of protein import in higher plants (Gagat et al., 2013; Schleiff & Soll, 2005).

1.2.2. Diversity and functional specificities of TOC GTPase receptors

Members of Toc159 family are characterized by three distinct domains: M- (membrane anchoring) domain, G- (GTP-binding) domain, and a highly acidic, intrinsically disordered A-domain (Fig. 1). There are four homologs in *Arabidopsis thaliana*: atToc159, -132, -120 and -90. While they share high similarity in their G- and M-domains, they largely differ in length and sequence at their A-domains (Hiltbrunner et al., 2001; Jackson-Constan & Keegstra, 2001). Toc34 proteins are smaller, membrane-anchored GTPases. In pea, only one member has been detected so far while two isoforms of Toc34 (atToc34 and atToc33) have been identified in Arabidopsis (Hust & Gutensohn, 2006; Jarvis et al., 1998).

Genetics and biochemical studies have supported the idea that various combinations of the different Toc GTPase isoforms lead to a diversity of complexes displaying differential selectivity for preprotein recognition and translocation (Constan et al., 2004; Ivanova et al., 2004; Kubis et al., 2003; 2004). Co-immunoprecipitation experiments performed by Ivanova and collaborators demonstrated that atToc159 preferentially associates with atToc33, while atToc120 and/or atToc132 preferentially form a complex together with atToc34 (Ivanova et al., 2004). Interestingly, the *toc34 (ppi3)* knock out mutant has no visible defect, while the *toc33 (ppi1)* mutant displays a pale green phenotype with a chloroplast biogenesis defect similar (although much less severe) than the *toc159* mutant phenotype (*ppi2*), supporting the proposition that these latter two receptor isoforms function in the same complex and preprotein import pathway (Bauer et al., 2000; Constan et al., 2004; Jarvis et al., 1998; Kubis et al., 2003; 2004).

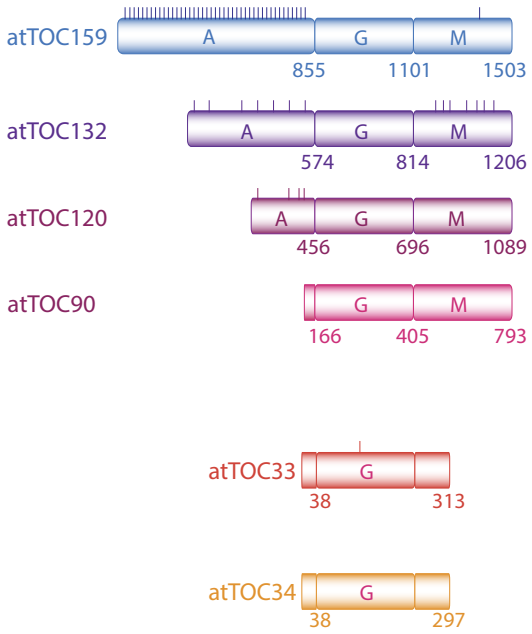


Figure 1. The TOC complex GTPases.
The TOC complex includes two GTPase-receptors Toc159 and Toc34. Toc159 consists of a GTPase domain (G) flanked by a C-terminal membrane anchoring domain (M) and an acidic N-terminal region (A). In Arabidopsis four Toc159 and two Toc34 isoforms have been identified. Toc159 homologs differ primarily in their A-domain sequences and lengths. Experimentally identified phospho-serine and -threonine residues (PhosphAT 4.0) are schematically indicated by short vertical lines.

Several lines of evidence indicate a potential functional overlap of the two Toc34 members: the strong sequence similarity : 65% (Jarvis et al., 1998); the fact that a minor fraction of atToc33 was co-immunoprecipated with Toc120/132, and atToc34 was detected with atToc159 (Ivanova et al., 2004); the embryo lethal phenotype of *toc33/toc34* double mutants and, most importantly, the ability of atToc34 to complement *ppi1* phenotype (Constan et al., 2004; Jarvis et al., 1998; Kubis et al., 2003). Transgenic complementation studies also indicated the potential functional overlap of atToc120 and atToc132 (Ivanova et al., 2004; Kubis et al., 2004) and, to a limited extent, for atToc159 and atToc90 (Infanger et al., 2011), however no functional overlap exists between these two subgroups (atToc120/132 vs atToc159/atToc90 (Ivanova et al., 2004; Kubis et al., 2004). While the two Toc34 homologs are mutually exchangeable, the same is only partially true for the Toc159 homologs, suggesting that preprotein selectivity of TOC complexes is mostly conferred by the identity of the Toc159 isoforms.

The classification of the client proteins of each isoform has been attempted. Because of the albino phenotype of *ppi2*, it has been proposed that Toc159 primarily facilitates the import of photosynthesis-associated preproteins. On the other hand, Toc132 or Toc120

being present predominantly in roots could facilitate that of constitutive (housekeeping) preproteins (Ivanova et al., 2004; Kubis et al., 2003; 2004; Smith, 2004). *In vitro* import assays using a selection of substrates support this model (Inoue et al., 2010; Smith, 2004). However, the albino phenotype of the *ppi2* mutant was shown to result not only from a defect in the import of a set of chloroplast proteins, but also from the transcriptional downregulation of a specific set of nuclear genes associated with photosynthesis (Bauer et al., 2000; Kakizaki et al., 2009). This effect is commonly referred to as retrograde signaling, and pleiotropically affects albino and pale green mutants across the board. The interference of retrograde signaling with preprotein import in *ppi* mutants has blurred the identification of the specific substrates of each of the receptor isoforms. Comparative analysis of *ppi2* mutant proteome and transcriptome demonstrated that certain photosynthesis-associated proteins accumulated normally in plastids even in the absence of atToc159, whereas accumulation of some house-keeping proteins were strongly diminished despite their mRNA expression levels being similar to the wild type (Bischof et al., 2011). Furthermore, the results of a yeast two hybrid screen used to identify the preferred Toc receptor of a variety of preproteins supported to the finding of Bischof and colleagues (Dutta et al., 2014). Together these studies affirmed that Toc GTPases, especially the Toc159 homologs, confer specificity to plastid preprotein import. However, specificity is not likely to be based on the photosynthetic or housekeeping nature of a preprotein. This is a move away from the overly simplistic paradigm of "photosynthesis-associated" and "house-keeping" specificities towards a more differentiated model that reflects complex and varying plastid preprotein requirements during development and under environmental influence. Therefore, Toc client protein classification will need to be rethought along these lines. One hypothesis is that the combination of preprotein specificities of plastid resident Toc receptors reflects the tissue or cell specific preprotein accumulation patterns that are specific to a particular plastid type.

As mentioned above Toc159 homologs diverge the most at their A-domains, suggesting a key role in their functional specialization. In domain swapping experiments, Inoue et al., replaced the A-domain of atToc132 by that of atToc159. Expression of this construct partially restored chlorophyll accumulation in the *toc159* null mutant (*ppi2*), while no complementation was observed using a construct encoding atToc132 without an A-domain. These data elegantly demonstrated that the functional specialization relies at least partially on intrinsic properties of the A-domain (Inoue et al., 2010). In agreement with this, it was observed that removal of the A-domains of atToc159 and atToc132 reduced the binding selectivity of these isoforms (Dutta et al., 2014; Inoue et al., 2010; Smith, 2004). Apparently, the A-domain does not directly interact with preproteins but may act as a filter enhancing the affinity for subsets of proteins and reducing the affinity for others (Dutta et al., 2014). Preprotein binding to Toc159 has

been shown earlier to occur at the G-domain (Smith, 2004). Thus it seems likely that the A-domain influences the G-domain by, for instance, positively or negatively modulating access of a preprotein according to its nature. Finally, the lack of complementation of *ppi2* by atToc132 lacking an A-domain (Inoue et al., 2010) as well as the recent work of Smith and colleagues using a yeast two hybrid system to study the preprotein-Toc159 receptor isoforms affinity (Dutta et al., 2014) indicates that a degree of specificity is conferred by the G-domain itself.

1.2.3. Diversity and complexity of the transit peptides

Inherently, recognizable specificity features would need to be encoded in the plastid transit peptides. One general consideration regarding the transit peptides is that no consensus can be defined, even when considering the structure at the three dimensional level (Bruce, 2001; Heijne & Nishikawa, 1991). Plastid transit peptides largely vary in length from an average of 50 up to 146 amino acids (Li & Teng, 2013a). There are some features shared with mitochondrial targeting peptides such as the overrepresentation of serine and threonine residues that may explain the targeting of plastid transit peptide containing proteins to mitochondria when expressed in heterologous animal systems(Zhang & Glaser, 2002). No further similarities between plastid and mitochondrial targeting sequences have been identified, and other levels of specificity might exist and enable plant cells to discriminate and accurately sort the two types of organellar proteins. Interestingly, an estimated thirty percent of chloroplast localized proteins do not have a canonical transit peptide (Ferro et al., 2003; Jarvis, 2008b; Kleffmann et al., 2004; 2007; Leister, 2003). A recent study in pea indicated that this may be an overestimation that results from a slightly inaccurate algorithm that does not take into account the whole diversity of features of plastid transit peptides (Chang et al., 2014).

The diversity of transit peptides sequences might well be explained by the need to fine tune the import of specific subsets of proteins in agreement with plastid type and developmental stage. Toc159 binds preproteins via their N-terminal, transit peptides (Smith, 2004), so one might reasonably expect that the specificity determinants reside within this particular region. However, the determining sequence elements that confer selectivity to a Toc159 isoform have not yet been identified. They could consist of cryptic signals buried in motifs and multiple-motifs (Bionda et al., 2010; Chotewutmontri et al., 2012; Lee et al., 2009a). For example Lee and colleagues revealed that Toc159-dependent import can be mediated by multiple independent motifs, one that consists in a stretch of serine residues located in first 12 amino acid of the N-terminal region of preRBCS (pSSU), and one located in the C-terminal part of the transit peptide sequence (Lee et al., 2009a). In a recent review, Li and Teng (Li & Teng, 2013) analysed such motifs and their relation with binding sites for various proteins involved in preprotein import. The authors then attributed the preproteins to distinct

subgroups based on patterns of sequence motifs in combination with their capacity to be targeted and bind to the protein translocon at the chloroplast outer envelope. Though only a limited number of preproteins were taken into account in these analyses, they clearly indicated that complexity of transit peptide design plays a key role in import selectivity.

1.3. Regulation of Toc components

1.3.1. Expression pattern

Regulation of TOC complex activity occurs at several levels. Overall the accumulation levels of Toc components throughout development appear to reflect the total import activity, i.e. a highest level of expression for the different components is observed in young, developing tissue, as compared to mature organs (Ivanova et al., 2004; Jarvis et al., 1998; Kubis et al., 2004; Kubis et al., 2003; Yu & Li, 2001). As an exception, Toc90 appeared to be uniformly expressed throughout development (Infanger et al., 2011; Kubis et al., 2003). Specific patterns were revealed when comparing the expression levels of the different Toc receptors isoforms in different organs and/or different plastid types, and usually correlated with corresponding mutant phenotypes in Arabidopsis (Bauer et al., 2000; Gutensohn et al., 2000; Jarvis et al., 1998; Kubis, 2004; Yan et al., 2014).

atToc159 and atToc33 are the most highly expressed members of their respective families and both mutants displayed the most severe visible phenotype when compared to other single mutants (Jarvis et al., 1998; Kubis et al., 2004). Furthermore, defects of plastid development in the corresponding mutants follow the expression pattern of the corresponding gene: highly regulated expression is observed for atToc159 and atToc33, with a higher expression occurring in photosynthetic tissues, when compared to other family members. Accordingly single mutants of these genes are specifically affected in plastid type present in those tissues, i.e. the chloroplast and its precursor, the etioplast (Bauer et al., 2000; Jarvis et al., 1998). By the same token, the higher expression of atToc120 and atToc132 in roots correlates with a severe defect of root plastid development in the corresponding double mutant (Kubis et al., 2004). Similarly the mutant phenotype of atToc34, which is expressed more highly in roots, retains normal plastid development but displays reduced root length (Constan et al., 2004; Gutensohn et al., 2000). Thus, selectivity of import into plastids can be modulated at least in part by transcriptional regulation of Toc components in accordance with plant tissue and/or growth conditions (light conditions in the case of Toc159).

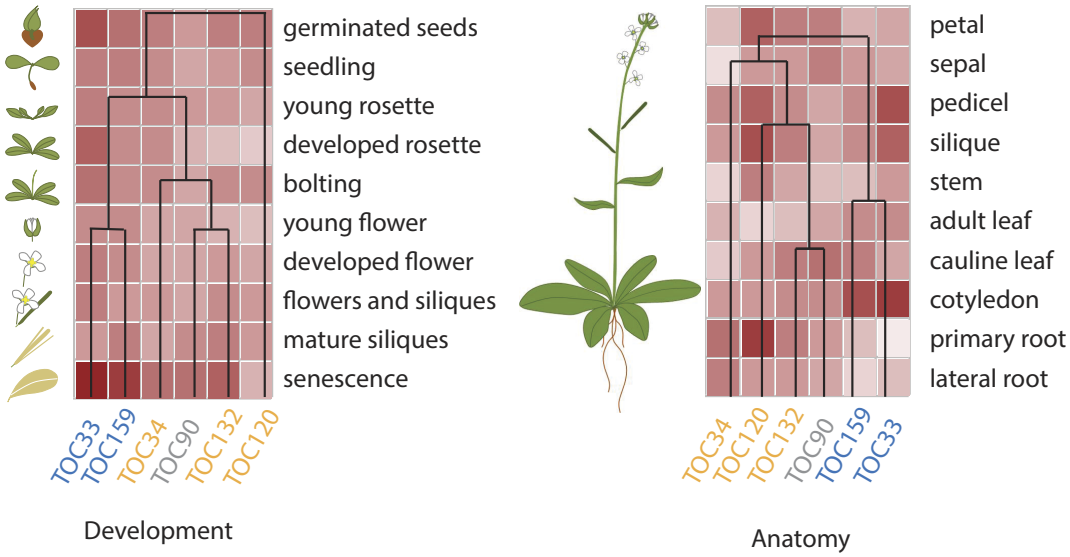


Figure 2. Hierarchial Cluster analysis of Arabidopsis Toc receptor expression.
TOC complexes consist of the assemblies of two different receptors from two separate GTPase families, Toc159/-132/-120/-90 and Toc33/-34, respectively, together with the Toc75 channel. Biochemical and genetic evidence have shown that atToc159 preferentially associates with atToc33 whereas atToc132 and 120 preferentially associate with atToc34. These specific associations are reflected by co-regulation of the Toc receptors isoforms. Data were extracted from Genevestigator database (Nebion), using the Hierarchical Cluster analysis tool, with "Development" or "Anatomy" specific selections for left- and right-hand panels, respectively

Expression profiles of the different Toc members suggest that the receptors acting together in a specific complex are co-regulated at the transcriptional levels. Interestingly, hierarchical cluster analysis indicates that this co-regulation extends to a large variety of conditions (Fig. 2) and suggests that common *cis* and *trans* regulatory elements could regulate associated Toc receptors. In support of this idea, the CIA2 transcription factor was found to co-modulate atToc33 and atToc75 expression specifically in leaves (Sun et al., 2001; 2009). However, the identity of other transcription factors responsible for the differential expression of Toc members has been poorly investigated so far and further experimentation will be necessary to reveal the molecular mechanisms underlying the regulation of Toc gene expression.

1.3.2. Post translational modifications

Differential regulation of Toc components also occurs at the post-translational levels (Fig. 3). It is interesting to note that the *ppi2* mutant can be complemented by expression of atToc159 under the constitutive 35S promoter indicating that transcriptional regulation can be bypassed at least under laboratory conditions (Agne et al., 2009; Kubis et al., 2004).

1.3.2.1 Phosphorylation

Several studies have shown that Toc receptors are phosphorylated. Phosphorylation has been reported for pea Toc34 and its ortholog atToc33 (Ser113 and S181, respectively), while it was not detected for atToc34 (Fulgosi & Soll, 2002; Jelic et al., 2003; 2002; Sveshnikova et al., 2000). Differential phosphorylation could therefore represent a regulatory mechanism conferring specificity to the two different members of Arabidopsis Toc34 family.

In vitro studies indicated that phosphorylation has a negative effect on GTP and preprotein binding to psToc34 and atToc33 (Jelic et al., 2003; Sveshnikova et al., 2000). Furthermore, *in vitro* and *in vivo* data showed that phosphorylation/phosphomimicking at atToc33 and phosphorylation of psToc34 negatively influenced TOC complex integrity (Oreb et al., 2008). Hypotheses for the underlying molecular mechanisms have been put forward. Since GTPase activity may be required for G-domain-mediated association of Toc159 and Toc34 (Smith, 2002; Wallas et al., 2003), phosphorylation may indirectly prevent homo- as well as heterodimerization because of a negative effect on GTP-binding. More directly the bulky, negatively charged phosphate group could inhibit the binding to a preprotein or to Toc159. However, this latter hypothesis may be valid for Arabidopsis, but not for pea since the phosphorylation site is distant from the dimerization interface (Oreb et al., 2008). In summary, the available data suggest the phosphorylation of psToc34 and atToc33 have a dual function, regulating both TOC complex assembly and subsequent substrate binding.

The physiological relevance and the signals triggering this specific phosphorylation are still not clearly defined. Data obtained from Arabidopsis transgenic lines expressing phosphomicking variants of atToc33 confirmed that phosphorylation at S181 can inhibit atToc33 activity in young Arabidopsis seedlings but not later during development (Aronsson et al., 2006; Oreb et al., 2007). Indeed, phosphomimick variants resemble the *ppi1* mutant regarding a number of phenotypic traits in 5 day-old Arabidopsis seedlings (chlorophyll accumulation, chloroplast ultrastructure and photosynthetic activity). However, since the non-phosphorylatable version behaved similarly to the WT, it was not possible to determine the conditions under which atToc33 is phosphorylated in *planta* (Aronsson et al., 2006; Oreb et al., 2007). We speculate that phosphorylation might represent a means to quickly down-regulate preprotein import *via* atToc33 containing Toc complexes, for example in mature plastids where protein demand is low. Moreover and since atToc33 can be phosphorylated but not atToc34, this post-translational regulation may affect the selectivity aspect of preprotein import regulation.

One additional phosphorylation site has been experimentally identified in both atToc33 and -34 (data provided by PhosphAT (Durek et al., 2010).

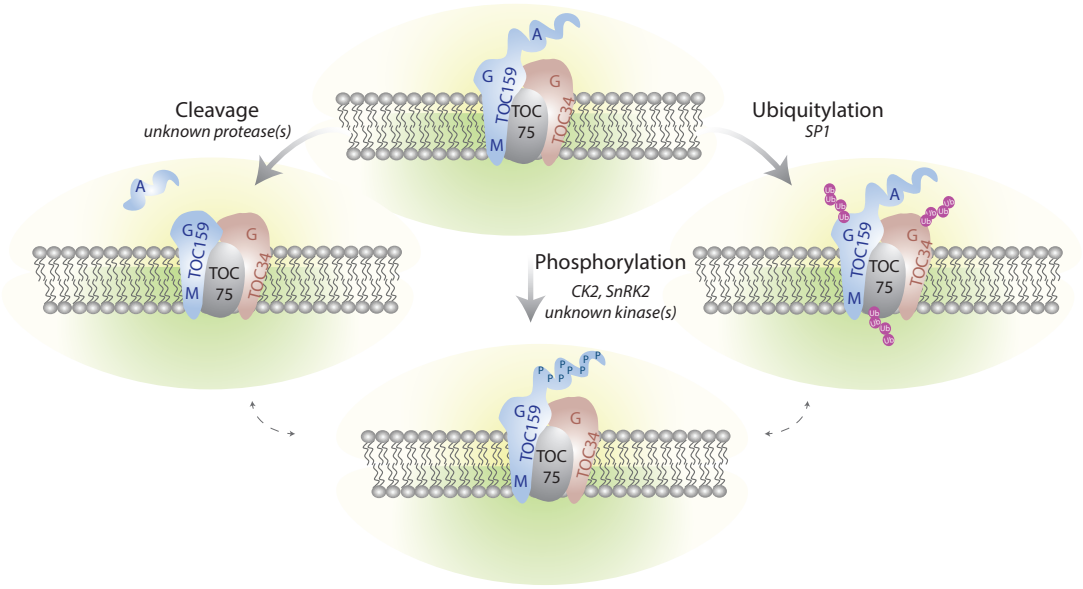


Figure 3. The TOC complex is targeted by multiple post-translational modifications
Phosphorylation of Toc159 and Toc34 at the G-domain may regulate homo and heterodimerization of the Toc receptors as well as their interaction with preproteins. The A-domain of Toc159 is hyperphosphorylated and can be released from the rest of the protein by proteolysis. The functional significance of these A-domain modifications are unclear but they may modulate the selectivity of the receptors for their client preproteins. All the Toc components are subject to ubiquitylation. Ubiquitylation may serve as a signal for proteasome-mediated degradation and pave the way for remodeling of TOC complex composition during plastid differentiation or environmental adaptation. The signaling pathways as well as the environmental and/or developmental factors triggering these post-translational modifications remain poorly described. Whether crosstalk between the different types of regulation exists is also not yet known.

It maps to a conserved Tyrosine residue of the G-domain. Additional studies will be required to validate and determine the regulatory effect of this specific phosphorylation.

Finally, the identity of Toc33/Toc34 kinase(s) still remain(s) mysterious. Some clues stemming from pea suggest that psToc34 is phosphorylated by an ATP-dependent, 98 kDa kinase residing at the outer envelope membrane (Fulgosi & Soll, 2002). However, the amino acid sequence information is not sufficient to molecularly identify the potential kinase in pea or its homolog in Arabidopsis.

The Toc159 receptors are also targets of phosphorylation. First evidence of phosphorylation of Toc159 came from *in vitro* studies using outer envelopes isolated from pea chloroplasts, showing that both full length Toc159 and its natural 86 kDa fragment could be phosphorylated (Fulgosi & Soll, 2002). Phosphorylation was

demonstrated for the G-domain of psToc159, reminiscent of Toc33/34 regulation (Oreb et al., 2008), however neither the precise site nor the regulatory function were further investigated. Large-scale phosphoproteomics projects revealed that Toc159 members in Arabidopsis are highly phosphorylated at the acidic A-domain (Agne et al., 2010; Durek et al., 2010). In total, 43 sites have been mapped in atToc159, while far fewer were detected in the other three members. These lower numbers may be due to the shorter length of the atToc132 and atToc120 A-domains, the absence of such a domain in atToc90, or because lower protein accumulation levels when compared to atToc159 limit the detection by mass spectrometry. Nevertheless the identified phosphorylation sites do not map to matching positions in the different homologs, which confers an additional degree of divergence to the A-domain.

The functional relevance of A-domain phosphorylation has been poorly documented so far. The dispensable nature of the A-domain suggests that phosphorylation either plays a minor role altogether, or possibly an important regulatory role under specific conditions (Agne et al., 2009; Hiltbrunner, 2001; Inoue et al., 2010). The A-domain behaves as an intrinsically disordered protein, which is often linked to multiple and transient protein-protein interactions (Richardson et al., 2009). Therefore phosphorylation of this domain could modulate interactions of Toc159 with other Toc components but also with specific sets of client preproteins. In addition, a selective autoinhibitory function of the A-domain under specific conditions may be envisaged that may be alleviated by phosphorylation or proteolytic removal.

Recently a link between ABA signaling and phosphorylation of Toc159 family members in Arabidopsis has been established (Wang et al., 2013). Upon ABA treatment atToc159 was phosphorylated at Thr692. atToc120 and atToc132 phosphopeptides accumulation was also enhanced by ABA treatment. These data together with the fact that a mutant deficient in ABA synthesis is affected in pre-protein import and early plant development suggest a close link between ABA signaling and chloroplast protein import regulation via Toc159 A-domain phosphorylation (Zhong et al., 2010). Whether ABA dependent phosphorylation plays a role in preprotein recognition, impacts TOC159 complex assembly, or acts at the level of the translocation process will be interesting questions to be addressed in the future.

Several classes of kinases may mediate phosphorylation of Toc159 homologs. Motif analysis suggests that a large fraction of atToc159 phosphorylation sites represent potential cytosolic casein kinase 2 (CK2) targets and this was validated biochemically by *in vitro* phosphorylation experiments (Agne et al., 2010). Recently it has been shown that ABA dependent phosphorylation of atToc159 at Thr692 was decreased in a triple mutant *snrk2.2/2.3/2.6* that is nearly insensitive to ABA treatment (Wang et al., 2013). In addition SnRK2.6 phosphorylated recombinant atToc159 *in vitro*. Thus SnRK2.6

represents a potential kinase of atToc159 at Thr692. On the contrary, atToc120 and atToc132 phosphorylation upon ABA treatment was detected only in the triple mutant *snrk2.2/2.3/2.6*, indicating the involvement of another ABA regulated kinase. Indeed ABA signaling is mediated by multiple kinases of the SnRK family but also of the MAPK kinase family (Danquah et al., 2014). The phosphorylation status of Toc159 members could therefore be regulated antagonistically by ABA signaling via the action of different classes of kinases and could represent a way to switch between Toc132/Toc120 to Toc159 specific import depending on environmental as well as developmental conditions and consequent plastid preprotein requirements. Finally, it has been proposed that psToc159 is a target of a 70kDa kinase located at the outer envelope of the pea chloroplast (Fulgosi & Soll, 2002) but so far no study has reported on the identification of a putative homolog in Arabidopsis.

In conclusion phosphorylation of the Toc159 and Toc34 receptors potentially regulates protein import at different levels: it may impact the import rate by regulating the affinity toward client preproteins, or affect the composition of the TOC complex by modulating the interaction between Toc receptors and consequently change the selectivity of plastid protein import. The involvement of ABA signaling in this regulation indicates that phosphorylation of Toc components can modulate the import activity in response to developmental signals for example during germination or subsequent post-germinative processes, or in response to abiotic stress that require the tuning of the plastid proteome. Hormonal control of plastid development has been frequently reported, but the effects on import activity are still poorly documented.

Phosphorylation could also be part of a signaling cascade enabling subsequent additional post-translational modifications since cross talk between different post-translational modifications (PTM) is a common phenomenon in eukaryotic systems, and PTM other than phosphorylation have been described for the different Toc components (see below). The existence of numerous phosphorylation sites, especially in Toc159 families, suggests the participation of multiple kinases and corresponding signaling pathways probably acting in a network.

1.3.2.2. PTM other than phosphorylation

Toc159 was first identified as an 86kDa protein lacking the A-domain (Bölter et al., 1998; Hirsch et al., 1994; Kessler et al., 1994; Schnell et al., 1994). It is not clear whether proteolysis occurs only during chloroplast preparation or whether it is part of regulatory system acting on Toc159. It is not clear either if other Toc159 homologs are also substrates of proteolytic cleavage but the relative stability of the A-domain fragment of atToc159 favors controlled proteolysis (Agne et al., 2010). Therefore, a yet unknown protease may process Toc159 conditionally, leading to the removal of the A-domain and consequently altering the import selectivity. Interplay between

phosphorylation and cleavage has been demonstrated in other biological systems for example in the context of apoptosis (Dix et al., 2012). Investigation of the cross talk between these two PTM will certainly be an interesting aspect for future research.

Abundance of the different Toc members varies developmentally. Currently an important question is to understand how the TOC machinery is remodeled upon plastid development and plastid inter-conversion. As discussed above transcriptional regulation plays a role in modulation of Toc components expression depending on plant tissues and environmental conditions, while post-translational modifications may participate in the regulation of TOC complex assembly and activity. Recently a genetic study complemented by biochemical analyses revealed that Toc receptors as well as the Toc75 channel could be modified by ubiquitylation. Ubiquitylation required SP1, a chloroplast outer membrane localized E3 ubiquitin ligase (Ling et al., 2012). Enhanced accumulation of TOC proteins in *sp1* genetic background suggested that SP1 indeed participates in UPS-mediated degradation of Toc components. Phenotypic analyses indicated that this regulatory mechanism may play a role during plastid inter-conversion. However, how SP1 is regulated and functions selectively on the different Toc receptors has not been addressed so far. Again a possible interplay with phosphorylation regulation might be envisaged as phosphorylation can serve as either a positive or a negative signal for ubiquitylation (Hunter, 2007).

The functional importance of proteolytic removal of the Toc159A is poorly understood; mainly because the enzyme(s) catalyzing the cleavage is not yet identified. However, identifying the protease will only partially solve the problem. The functional relevance of the significant accumulation of Toc159A must be explored. At present the only available comparison is the similarity in the physical properties of Toc159A domain with some of the acidic ribosome subunits that are associated with translation machinery in the cytosol (Bautista-Santos & Zinker, 2014). So exploring the relevance of cytosolic abundant Toc159A with preprotein translation might help us understand its functions as a soluble protein.

1.4. Aims of this study

Chloroplast biogenesis is a key step during the onset of photoautrophic growth. To accomplish this developing chloroplast imports around 2000 different preproteins from the cytosol. The import of many of these is mediated by the preprotein receptor Toc159 located at the surface of the organelle. Initially, Toc159 was identified as a 86kDa protein at the chloroplast outer envelope of pea (*Pisum sativum*) (Schnell et al., 1994; Waegemann & Soil, 1991). Genome sequencing in Arabidopsis lead to the identification of a 159kDa ortholog as well as three additional homologs Toc132, -120 and -90. Toc159, -132 and -120 have three distinct domains A (acidic), G (GTPase) and

M (membrane insertion) whereas Toc90 has only G- and M-domains (Bauer et al., 2000; Bölter et al., 1998). In Arabidopsis, Toc159 without the A-domain fully complemented the albino phenotype of the *ppi2* mutant plants lacking Toc159. This finding indicated that Toc159A domain has no essential function (Bölter et al., 1998; Hiltbrunner, 2001).

However, recently published data require us to reconsider the function of the acidic domain in full length Toc159. Several studies offer clues for the potential role of the acidic domain in preprotein selectivity for import either by Toc159 or its homologs (Ivanova et al., 2004; Kubis, 2004; Smith, 2004). Inoue and colleagues (Inoue et al., 2010) tested the affinity of the A-domains of Toc159 and Toc132 towards a few selected preproteins *in vitro* and came to the conclusion that the acidic domains of these two GTPases have distinct preprotein specificities which may help to avoid competition for import sites at the chloroplast surface. Based on yeast two hybrid experiment, Dutta and colleagues (Dutta et al., 2014) reached the conclusion that the acidic domains might act by inhibiting binding of certain preproteins to the G domain of the corresponding GTPases. Our group reported proteolytically-removed Toc159A in the cytosol and at the chloroplast existing in a hyperphosphorylated state (Agne et al., 2010; Agne & Kessler, 2010) These finding suggest that both proteolysis and phosphorylation may contribute to regulating preprotein specificity at the level of the A-domain.

While these reports have significantly improved our understanding of preprotein recognition by Toc159 GTPases, they have also raised additional questions. These concern the molecular mechanisms and the nature of additional, unknown components (for instance kinases) that impart preprotein specificity to the acidic domains.

This study aims to discover the functions of Toc159A domain *in vivo* and *in vitro*
The objectives of this study are

1. Localisation of the A-domain by the expression of the fluorescently-tagged A-domain *in vivo* and proteolytic resistance of the A-domain *in vitro*. This knowledge is essential to understand the function of the protein.

2. Identification of new interacting proteins, such as kinases, that may elucidate the functions of Toc159A. *In vivo* protein-protein interactions will be studied by tandem affinity purification (TAP) using a TAP-tagged version of the A-domain.

3. *In vitro* experimentation will be carried out to test hypotheses regarding the molecular mechanism of the A-domain that contributes to preprotein selectivity. These include direct binding of preproteins and regulation of translation activity of the ribosome by the A-domain.

2. Results

2.1.Dual localisation of Toc159A

In a previous study Toc159A was identified both as part of and removed from the full length protein. Toc159A separated from the full length protein was present both at the chloroplast envelope as well as in the cytosol as a soluble and extensively phosphorylated protein (Agne et al., 2010) Since separated Toc159A appeared stable, we hypothesized that this fragment may have some kind of activity. The confirmation of the dual localisation of Toc159A, is therefore an important step to discover its function.

2.1.1. Molecular cloning of T-DNA vector for in planta expression and characterization of a Toc159A-YFP fusion protein

Since my interest was to further analyze the dual localisation of Toc159A, I engineered a gateway-based plant expression vector to express a Toc159A-YFP fusion protein under the constitutive, cauliflower 35S promoter (pEG101-3HA-Toc159A-YFP).

Gateway cloning was performed as described (see 4.2.1.4). Colonies containing the clone pEG101-3HA-Toc159A-YFP were identified by colony PCR (Fig.4a). The PCR reaction relied on one primer within the vector and another one inside the gene designed to give a product of 0.75 kb. Plasmid DNA was extracted from positive colonies and the correct sequence of on the plasmid preparation was confirmed by sequencing. (see appendix 1). I used this verified pEG101-3HA-Toc159A-YFP plasmid to transform protoplasts as described (see 4.2.6.2) and wild type Arabidopsis plants (see. 4.2.6.3). Transformed Arabidopsis plants were selected on phosphinothricin-containing medium (image inset Fig.4b) and visually selected for the presence of YFP fluorescence under a *Nikon* fluorescence binocular microscope equipped with a GFP filter. The majority of the phosphinothricin-selected plants were fluorescent (Fig.4b). Western blotting against the HA tag was used to confirm the expression of *pEG101-3HA-TOC159A-YFP:WT* in Arabidopsis (Fig.4c). Two out of three tested plants expressed 3HA-Toc159A-YFP having a relative mass of over 116kDa.

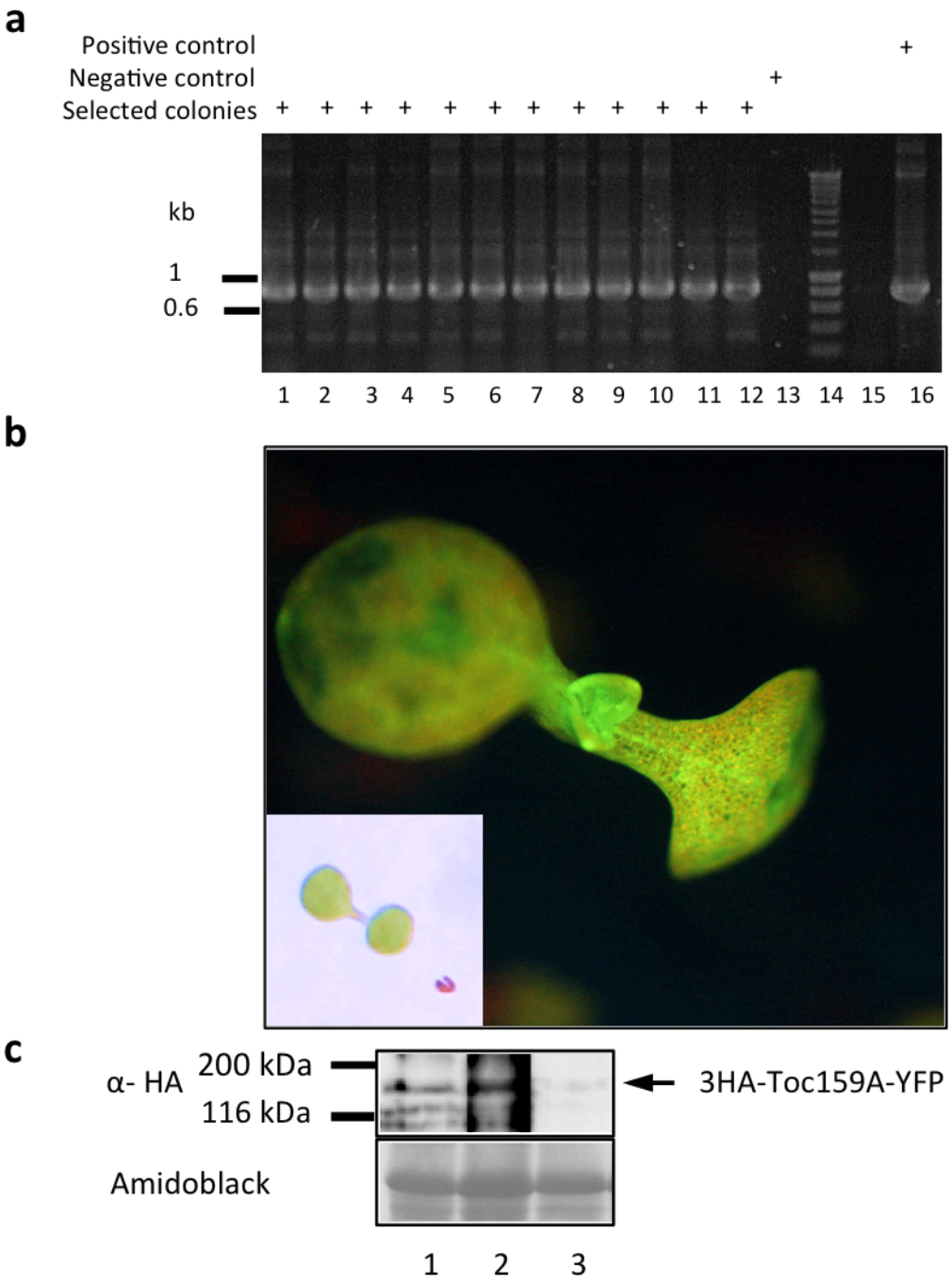


Figure 4. Molecular cloning of T-DNA vector for in planta expression and characterization of a Toc159A-YFP fusion protein

a) E. coli cells were chemically transformed with the product of Gateway cloning reaction of the *pEG101-3HA-TOC159A-YFP* vector. Antibiotic resistant colonies were screened by colony PCR. b) Wild type Arabidopsis plants transformed with *pEG101-3HA-TOC159A-YFP*, transformed plants were selected on Murashige-Skoog medium containing phosphinothricin (inset picture phenotypically they resemble the wild type plants grown on MS medium) and then analysed using a fluorescence binocular microscope using a GFP filter (emission 483 - 510nm). c) Transgenic *pEG101-3HA-TOC159A-YFP:WT* plants were tested for protein expression by Western blot using an antibody against the HA-tag.

2.1.2. Toc159A *in vivo* localisation using *3HA-TOC159A-YFP*

2.1.2.1. pEG101-3HA-TOC159A-YFP:WT protoplast transient expression

The *pEG101-3HA-TOC159A-YFP* construct was used to transfect wild type protoplasts for transient protein expression. Expression and localisation of the fusion protein was observed using a LEICA TCS SP2 confocal microscope and excitation wavelength of 514 nm (Fig. 5, panels A-F).

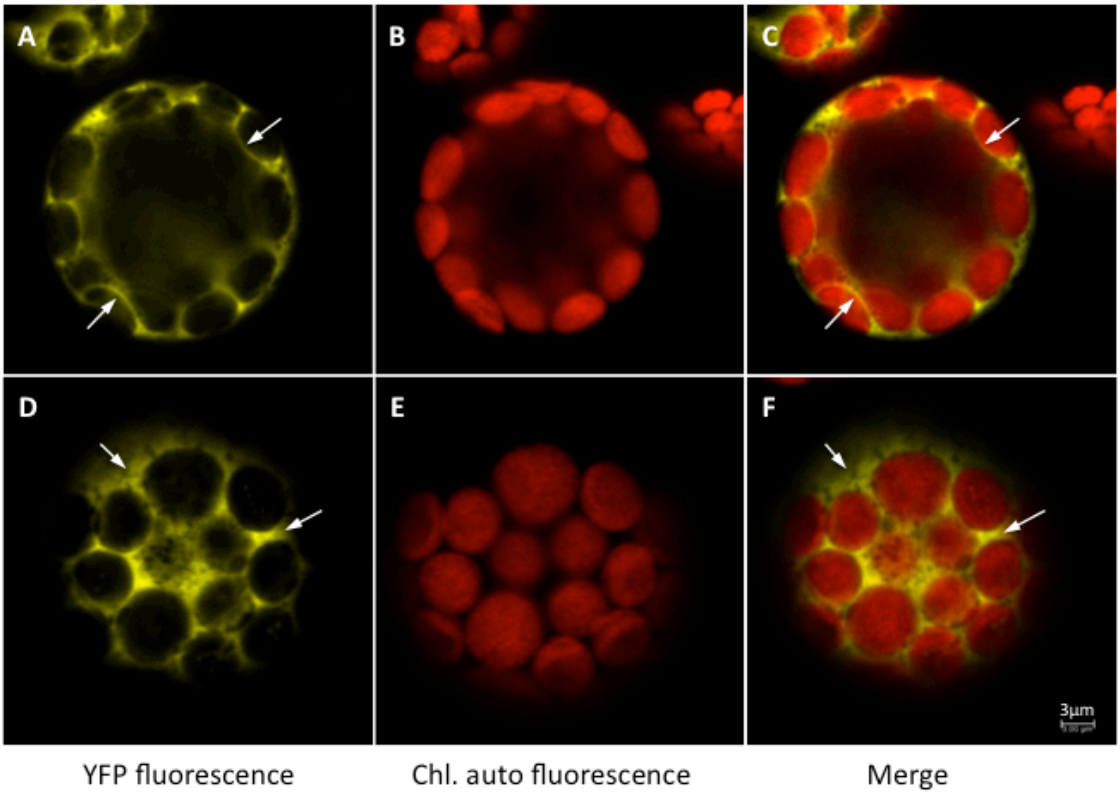


Figure 5. *In vivo* localisation of Toc159A in protoplasts.
pEG101-3HA-TOC159A-YFP was transiently expressed in wild type (Ws) Arabidopsis protoplasts. Transformed protoplasts were observed by confocal microscopy. (A-C) shows the 3HA-Toc159A-YFP fluorescent signal lining the chloroplast surface. Images (D-F) clearly show the predominant cytosolic presence of the same soluble Toc159A-YFP. This points to the ability of cytosolic Toc159A to associate with the chloroplast outer envelope when ectopically expressed as a soluble, separated domain.

Bright YFP fluorescence was observed in the open spaces between chloroplasts indicating cytosolic localization of overexpressed 3HA-Toc159A-YFP. In addition, bright YFP fluorescence was also observed in a thin line tightly surrounding the

chloroplast (Fig. 5, panels A-C) Such localisation patterns were previously observed in the case of chloroplast outer envelope proteins (Hiltbrunner et al., 2001b).

2.1.2.2. pEG101-3HA-TOC159A-YFP:WT Arabidopsis stable transformants

We examined *pEG101-3HA-TOC159A-YFP:WT* Arabidopsis stable transformants under the confocal microscope, and we could observe the similar pattern of fluorescent ring around chloroplasts besides the cytosolic presence (Fig. 6, panels A-C). This result was supported by the preliminary observation made in the previous study where the over expressed Toc159A-TAP was found in the chloroplast/ insoluble membrane pellet fraction (Agne et al., 2010). Though the constructs, plants and methods used in this study were completely different from the one used in the earlier report, the results were converging to confirm that indeed Toc159A might have the capacity of its own to be localized on chloroplast, however we were yet to find an explanation for this occurrence.

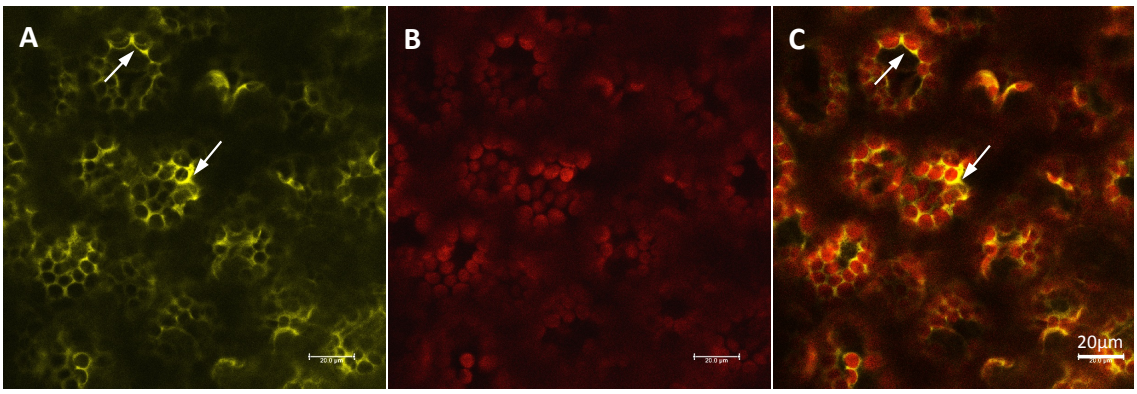


Figure 6. *In vivo* localisation of Toc159A in 3HA-TOC159A-YFP:WT plants
The dual localisation of 3HAToc159A-YFP in transgenic Arabidopsis expressing pEG101-3HA-TOC159A-YFP (A-C). C) A majority of the 3HA-Toc159A-YFP signal appears cytosolic but YFP-fluorescence was also observed at the chloroplast surface.

To summarize, endogenous Toc159A has previously been reported to be cleaved from full length Toc159 and appears as a stable, separate protein in cytosol as well as the chloroplast envelope (Agne et al., 2010). Here, we show that Toc159A is cytosolic and may also associate with chloroplasts when expressed as a soluble protein.

2.1.3. Toc159A thermolysin protease tolerance.

To further analyze the association of Toc159A with chloroplasts we addressed the nature of its association. To do so we used *TOC159A-TAP:WT* plants expressing Toc159A-TAP under the cauliflower mosaic virus 35S promoter (Agne et al., 2010).

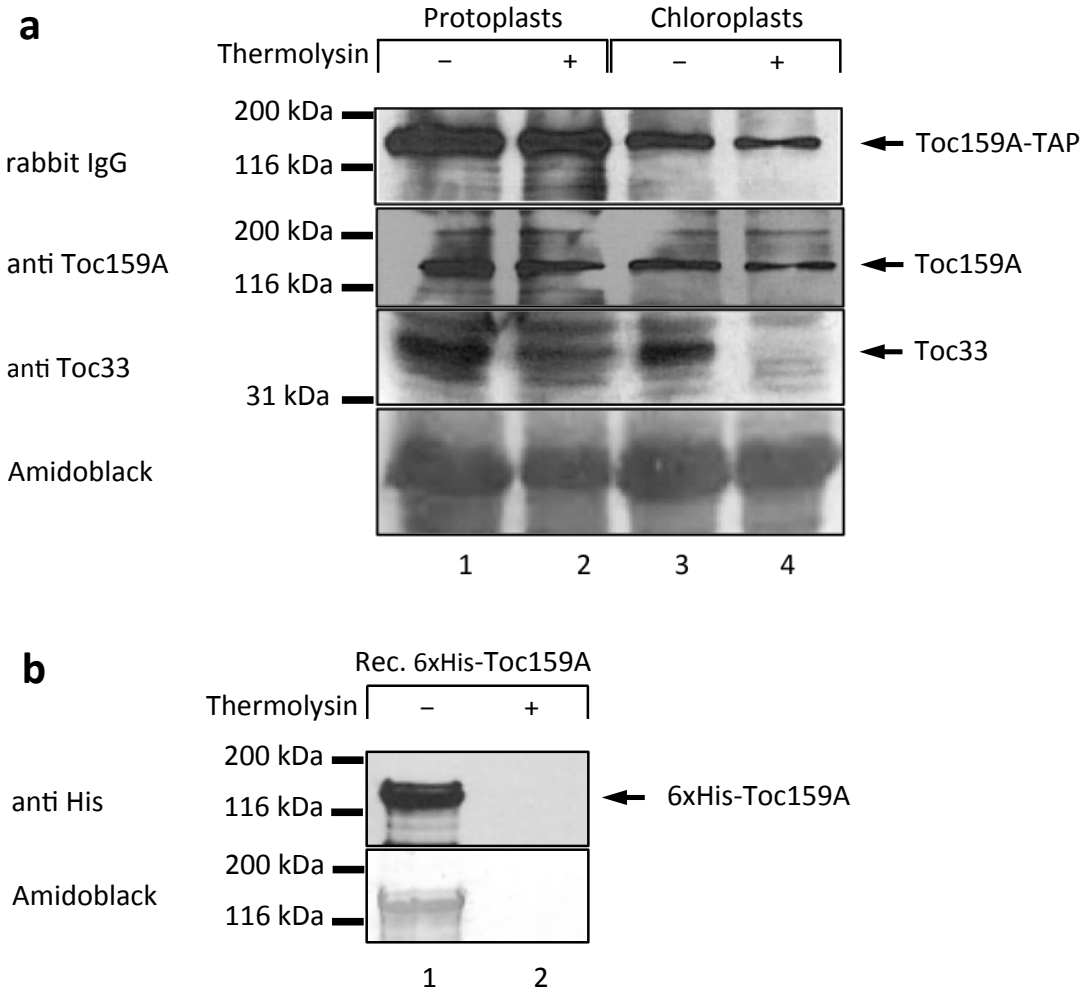


Figure 7. Biochemical localization of soluble Toc159A-TAP *in vivo*.
a) Protoplasts (lane 1, 2) and chloroplasts (lane 3, 4) from Arabidopsis seedlings expressing TOC159A-TAP under the 35S promoter in wild type background (Toc159A-TAP:WT) were incubated with thermolysin protease at a concentration of 100µg/mL. The immuno detection using Rabbit IgG reveals transgenic Toc159A-TAP, immuno detection using anti-Toc159A antibodies reveals both endogenous Toc159A and transgenic Toc159A-TAP. Western blotting using antibodies against Toc33 reveal its thermolysin sensitivity and was used as a positive control. Surprisingly, Toc159A-TAP was not thermolysin sensitive in chloroplasts and behaved as a membrane protected protein. b) To rule out that Toc159A is inherently thermolysin resistant 5µg of recombinant His_6-Toc159A (6xHis-Toc159A, lane 1) was subjected to thermolysin treatment (lane 2) at a concentration of 100µg/mL. Western blotting against the His_6 tag shows that 6xHis-Toc159A is completely degraded by thermolysin (lane 2).

The TAP (Tandem Affinity Purification)-tag contains IgG-binding motives separated from a calmodulin-binding peptide (CBP) by a tobacoo etch virus (TEV)-protease site. The TAP-tag can be detected in Western blotting experiments by anti-CBP antibodies or purified IgG recognizing only Toc159A-TAP but not the endogenous Toc159A. Endogenous Toc159A can be detected together with Toc159A-TAP using anti-

Toc159A antibodies. The TAP-tag also contains a TEV-site that can be used to release and elute a TAP-tagged protein upon affinity purification using a rabbit or human IgG column.

Intact protoplasts were prepared and subsequently used for the preparation of chloroplasts. Total protein was analysed by SDS-PAGE followed by Western blotting. Rabbit IgG was used to recognize the TAP tagged Toc159A, confirming A-domain localization at chloroplasts. Protoplasts as well as isolated chloroplasts were subjected to protease (thermolysin) treatment at a concentration of 100µg/mL, which digests proteins that are exposed at the chloroplast surface (Fig.7a). Under the conditions used no changes in Toc159A-TAP levels after thermolysin treatment of either protoplast or chloroplast preparations (compare Fig. 7a, lanes 1 and 2, and lanes 3 and 4) were observed.

The bottom panel of Fig. 7a, shows Amidoblack staining of the blot to indicate equivalent protein loading in all the lanes. To confirm that the thermolysin treatment was effective, we probed Toc33 by Western blotting, which is known to be a thermolysin sensitive protein at the chloroplast outer envelope. This result indicated that thermolysin was active during the course of this experiment.

To rule out Toc159A thermolysin insensitivity, a parallel experiment was carried out (Fig. 7b). Recombinant His_6-Toc159A, expressed in bacteria was purified using a two-step protocol on the ÄKTAPrime® system (see 4.2.3.1 and 2.3.1). Five micrograms of recombinant His_6-Toc159A protein were incubated in solution with thermolysin concentrations identical with those used for protoplasts and chloroplasts. Western blotting against the His_6-tag showed the presence of His_6-Toc159A in the untreated sample (lane 1), whereas thermolysin treatment completely digested the recombinant protein (lane 2). Amidoblack staining of the membrane (lower panel) confirms the finding.

The results confirm that Toc159A domain is not an inherently thermolysin resistant protein (Bölter et al., 1998). However the chloroplast-associated Toc159A-TAP was found to be thermolysin resistant; suggesting that it was protected in isolated chloroplasts. This may be due either to protection by outer chloroplast membrane or by the association of Toc159A with other proteins.

2.2. Cytosolic and chloroplast protein interactions of Toc159A

2.2.1. IgG pull-down of *in vivo* Toc159A-TAP complex

We hypothesize that Toc159A interacts with proteins that would assist Toc159A-YFP chloroplast localisation and provide thermolysin resistance to Toc159A-TAP. Hence I intended to purify such interacting proteins by tandem affinity purification of *TOC159A-TAP* from transgenic plants. In order to distinguish potential protein interactions in the cytosol from those at the chloroplast protoplasts were used because of their convenience to separate the two compartments. Crude cytosol fraction was isolated by breaking protoplasts and separating it from intact chloroplasts and broken membranes on a Percoll® gradient. Subsequently chloroplasts were detergent-solubilized to release membrane proteins (see. 4.2.7). Both fractions were used for affinity purification as described (see. 4.2.9.3). In parallel, an identical mock experiment was carried out using wild type plants not expressing Toc159A-TAP. Input (1%), flow through (1%FT) and TEV (tobacco etch virus) as well as SDS eluates (33%) were analyzed by SDS-PAGE followed by Western blotting.

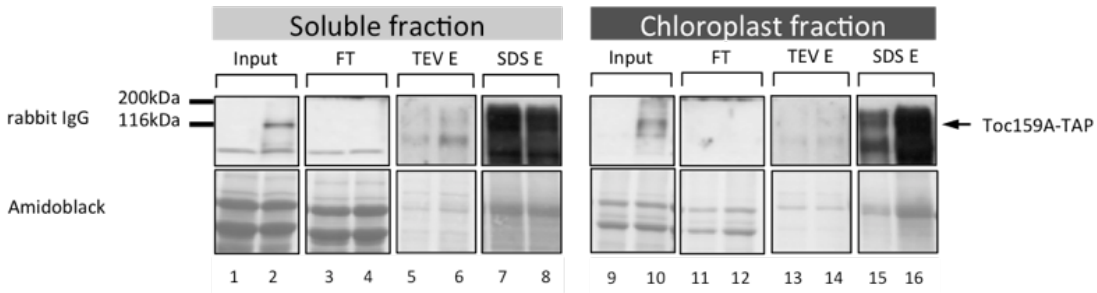


Figure 8. Tandem affinity purification of Toc159A-TAP complex.
Soluble fraction containing cytosol and chloroplasts solubilized in buffer containing 1% n-Dodecyl β-D-Maltopyranoside (DDM) from both wild type (lanes 1 & 9) and Toc159A-TAP:WT (lanes 2 & 10) plants were incubated with human IgG-sepharose. The flow through (FT) of both wild type (lanes 3 & 11) and Toc159A-TAP:WT (lanes 4 & 12) were collected. After thorough washing of the sepharose, bound proteins were first eluted with TEV protease and then with SDS. Proteins in the fractions were separated by SDS-PAGE and transferred to nitrocellulose. Proteins were stained by Amido black, Toc159A-TAP was detected by Western blotting using rabbit IgG. TEV Eluates (TEV E) of wild type (lanes 5 & 13) and of Toc159A-TAP:WT (lanes 6 & 14). 2% SDS Eluates (SDS E) of wild type (lanes 7 & 15) and of Toc159A-TAP:WT (lanes 8 & 16). The dark smear visible in both wild type and Toc159A-TAP:WT eluates is due to the leakage of IgG subunits from the IgG-sepharose column.

Toc159A-TAP was detected in the inputs (Fig. 8 lanes 2 & 10) derived from TOC159A-TAP:WT extracts but not in wild type extracts. Absence of this band in the unbound flow through fraction (lanes 5, 6, 11 & 12) suggested complete binding of Toc159A-TAP to the beads. However, we were unable to detect Toc159A-TAP in either the TEV

or the SDS eluates. Indeed, the band visible in the TEV eluates (lanes 5, 6, 13 & 14) was clearly migrating differently from Toc159A-TAP in the input, moreover it appeared both in Toc159A-TAP:WT (lanes 6 & 14) samples, as well as in the wild type mock experiment (lanes 5 & 13) therefore considered non-specific. The dark smear in the SDS-eluate can likely be attribute to IgG leakage from the affinity column (lanes 7, 8, 15 & 16).

2.2.2. Chemical crosslinking to increase Toc159A-TAP complex recovery.

Chemical crosslinking to increase the stability of protein-protein interactions has been used for a long time (Markwell & Fox, 1980; Phizicky & Fields, 1995). Lately the use of chemical crosslinkers in mass spectrometric studies has become common (Chavez et al., 2011). The advantage of these chemical reagents lies in the stabilization of crosslinked complexes towards various chemical and physical treatments during extraction and purification procedures. Amongst a large number of available crosslinking reagents, DSP (Dithiobis-[succinimidyl propionate]) was chosen for its many desirable features such as i) short spacer arm length of 12.0 Å (increases specificity) ii) small mass (facilitates membrane permeability) iii) reversible crosslinking (by reduction using DTT).

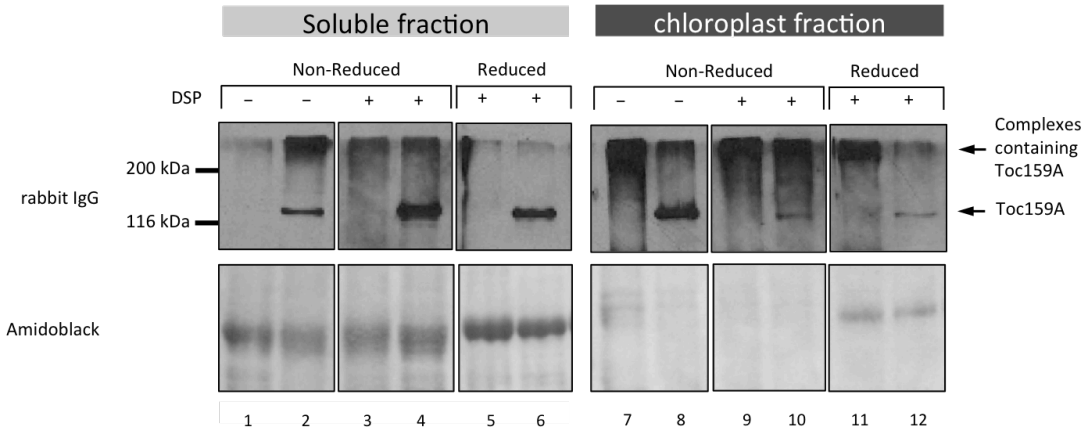


Figure 9. Reversible chemical crosslinking to stabilize Toc159A complexes.
Protoplasts from wild type (odd numbers) and TOC159A-TAP:WT (even numbers) were incubated in the presence or absence of 5mM DSP crosslinking reagent then separated into the soluble, cytosol-containing (Soluble fraction; lanes 1-6) and chloroplast fractions to identify Toc159A-containing complexes in these two compartments separately. The chloroplast pellet was solubilized with 1 % DDM (lanes 7-12). The fractions were analyzed both by reducing (Reduced, lanes1, 2, 3, 4 ,7, 8, 9 and 10) and non-reducing SDS-PAGE (Non-Reduced, lanes 5, 6, 11 and 12). Note that reduction reverses DSP crosslinking and thereby destabilizes complexes in SDS-PAGE gels.

Isolated protoplasts of wild type (Fig. 9, odd numbered lanes) and *TOC159A-TAP:WT* plants (Fig. 9, even numbered lanes) were treated with DSP and fractionated into soluble 'cytosolic' and membrane 'chloroplast' fractions. Chloroform:methanol

precipitated protein from these fractions were dissolved in non-reducing and reducing sample buffers, respectively, for SDS-PAGE and subsequent Western blot analysis. The samples dissolved in non-reducing sample buffer (Fig. 9, lanes 1-4, 7-10) were compared to those dissolved in reducing sample buffer (Fig. 9, lanes 5, 6, 11 & 12). The odd numbered samples from wild type protoplasts were also compared with even numbered *TOC159A-TAP:WT* samples. There is a pronounced increase in higher molecular mass signals, detected by rabbit IgG, in *TOC159A-TAP:WT* non-reduced samples both in soluble (lane 4) and membrane (lane 10) fractions. In the membrane fractions generally higher molecular mass unspecific aggregates appeared in *WT* samples regardless of the presence or absence of DSP and these aggregates interacted non-specifically with rabbit antibodies. However, potentially interesting DSP crosslinked complexes containing Toc159A-TAP were sensitive to DTT treatment (compare Fig. 9, lanes 4 with 6, 10 and 12), which lead to a visible decrease of higher mass signals in lane 12. It is therefore possible that these higher molecular mass signals correspond to *bona fide* Toc159A-TAP containing complexes.

2.2.3. Covalent crosslinking of protein interactions *in planta.*

Isolation of sufficient protoplasts for chemical crosslinking using DSP was a major limitation for mass spectrometric protein analysis. I therefore decided to attempt *in planta* crosslinking as an alternative.

Proteins of three week-old wild type and Toc-159A-TAP:WT plants were chemically crosslinked by immersion in a 5mM DSP solution and application of a vacuum for 5 min. The plants were washed with Tris-HCl buffer to quench the excess crosslinker on the surface then ground with mortar and pestle. The soluble fraction was separated from the membrane fraction by centrifugation. The samples were dissolved in non-reducing sample buffer prior to SDS-PAGE followed by Western blot analysis.
By Western blotting using rabbit IgG for detection (Fig. 10.), a dominant higher molecular mass band migrating above 200 kDa was detected in the Toc159-TAP:WT extract (lane 2) and may represent a Toc159A-complex.

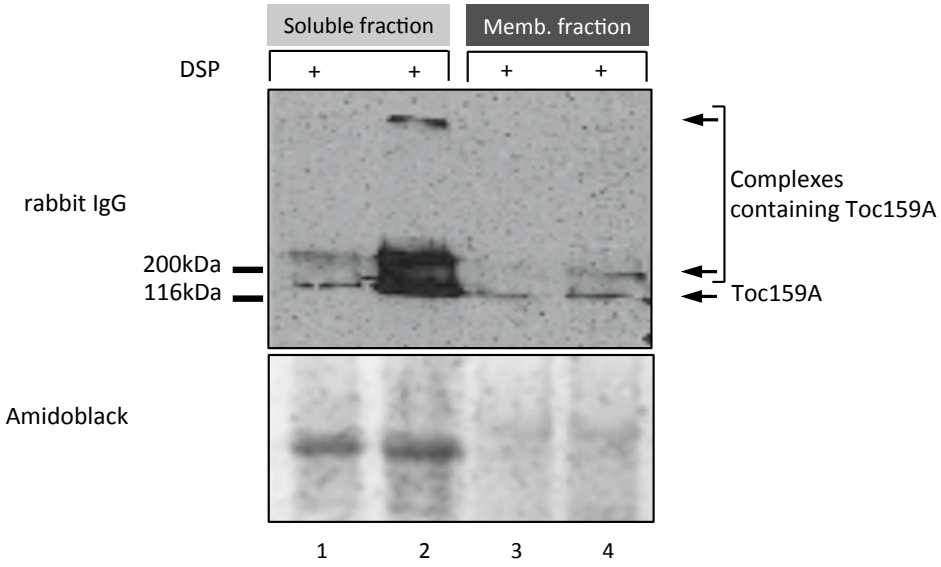


Figure 10. In planta chemical crosslinking by vacuum infiltration of DSP.
Both wild type (1 & 3) or Toc159A-TAP:WT (2 & 4) were subjected to DSP chemical crosslinking with aided by vacuum infiltration. Subsequently, plants were homogenized and separated into soluble fractions (1, 2) and membrane fractions from the total protein extract by centrifugation at 100,000g. The pellets were washed and then solubilized with 1 % DDM to prepare membrane fraction (Memb. fraction, lanes 3, 4).

2.2.4. Identification of Toc159A complexes by Blue native-Polyacrylamide Gel Electrophoresis (BN-PAGE)

After the attempts to identify and purify Toc159A complexes by immunoprecipitation and DSP crosslinking, I changed the strategy to identify the Toc159A complex by blue native polyacrylamide gel electrophoresis. For this purpose we used *TOC159A-TAP:WT* plants. Total Toc159A-TAP:WT protein was extracted with buffer containing 1% digitonin and prepared for BN-PAGE analysis. In parallel, Toc159A-TAP:WT extract was immunodepleted to remove Toc159A-TAP-containing complexes using anti-calmodulin-binding protein (anti-CBP) affinity chromatography specific to the TAP-tag (described in 4.2.5.3). In an additional experiment, anti-CBP antibodies were added to the Toc159A-TAP:WT extract. Purified recombinant (lane 4) was used as a control for the Toc159A monomer.

The Toc159A-TAP:WT extract (Fig. 11, lane 1) was run side-by-side with the immunodepleted sample (Fig. 11, lane 2), the one containing additional anti-CBP antibody (lane 3) and the purified, recombinant His_6-Toc159A (lane 4) on a BN-PAGE gel. The gel was then blotted to nitrocellulose and the blot probed using anti-Toc159A (recognizing endogenous Toc159A, Toc159A-TAP and purified recombinant His_6-Toc159A, Fig. 11, panel a) as well as anti-CBP (recognizing only Toc159A-TAP

(Fig.11, panel b). Purified recombinant His_6-Toc159A gave a major band at around 200 kDA (Fig. 11, lane 4). A band corresponding to this mass was detected in all samples suggesting the presence of monomeric Toc159A-TAP.

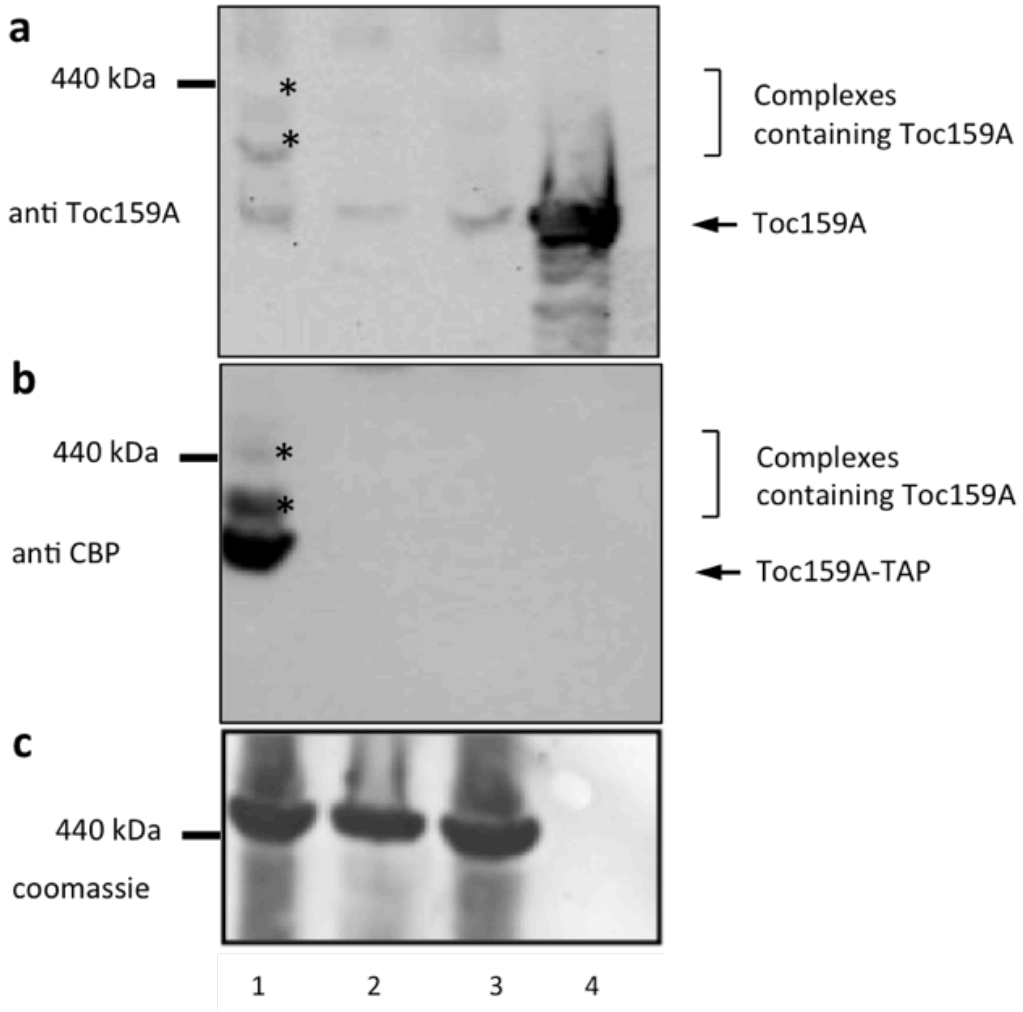


Figure 11 Identification of Toc159A-TAP complexes by BN-PAGE

To detect complexes containing Toc159A, total protein from Toc159A-TAP:WT plants was extracted in the presence of 1% digitonin and subjected to BN PAGE (lane 1). In lane 2, TAP-tagged proteins were immunodepleted from the extract prior to electrophoresis. Lane 3 is a lane 1 but rabbit IgG was added to the digitonin extract. Lane 4 contains purified recombinant His_6-Toc159A to indicate the migration of the Toc159A monomer. The BN-PAGE gel was tranferred to nitrocellulose and probed with anti-Toc159A antibody (top panel a) and anti-CBP (center panel b). The asterisks indicate potential higher molecular mass complexes. Bottom panel c) Coomassie stained nitrocellulose membrane of BN-PAGE after transfer to nitrocellulose. Note that the thick band migrating above 440 kDa corresponds to the RuBisCO complex.

In addition, several higher molecular weight bands were detected in lanes 1 of both panel a and b. These bands were not present in lanes 2 corresponding to the depleted extracts. This suggests that the bands represent complexes containing Toc159A-TAP.

The high molecular mass bands were also absent when the anti-CBP antibody (lane 3) was added to the digitonin extract. In this case (lane 3), the anti-CBP antibodies binding to Toc159A-TAP may have led to the formation of very high molecular mass aggregates that were unable to enter the gel and therefore may have escaped detection by Western blotting.

In summary, the BN-PAGE experiment confirms that Toc159A-TAP is present both as monomer and at least two higher molecular mass bands that probably represent protein complexes. However, to identify the proteins in those bands they need to be excised from the gel and subjected to tryptic digestion prior to mass spectrometric analyses. The close proximity of the putative Toc159A complexes to highly abundant RuBisCO (Fig. 11 panel c) on the gel interfered with mass spectrometry (Sacha Baginsky, personal information).

2.2.5. Formaldehyde crosslinking of *in vivo* protein interactions

BN-PAGE confirmed the existence of higher molecular weight complexes containing Toc159A-TAP, but the RuBisCO complex migrating very close to the Toc159A-TAP complexes makes it necessary to purify the complexes before protein identification. We learned from earlier experiments that the Toc159A-TAP complexes were unstable during immunopurification. Therefore, I still needed to use a crosslinker that would stabilize the complexes during the extraction and immunopurification procedure. Formaldehyde has been used as preservative of proteins in cells, tissues and entire organisms for a long time. Its tiny size and easy permeability into cells has contributed to its success. However, until recently the use of formaldehyde as a chemical crosslinker to study protein-protein interactions hasn't been explored (Sutherland et al., 2008).

Recent studies demonstrated that at optimal concentrations, formaldehyde may legitimately crosslink *in vivo* protein-protein interactions (Klockenbusch & Kast, 2010; Zhu et al., 2014). However, these studies were carried out using animal cell cultures and the conditions for formaldehyde crosslinking of protein-protein interactions in Arabidopsis still need to be established. To do so, TAP-TOC159:*ppi2* (an Arabidopsis line expressing full length, N-terminally TAP-tagged Toc159 in the *ppi2* background)

was used because the TOC159 complex has already been well characterized(Chen & Li, 2006).

TAP-Toc159:*ppi2* as well as NTAP:WT control seedlings were infiltrated with 0, 0.5, 2, 4, 6 and 10% formaldehyde solutions by applying a vacuum; after decanting the solution any remaining formaldehyde was quenched by the addition of glycine. Total protein extracted from the crosslinked plants was separated on a 4-16% gradient SDS-PAGE gel followed by Western blotting to analyze stable complexes. By Western blotting using rabbit IgG to detect the TAP tag (Fig.12) a number of high molecular mass bands were detected in TAP-Toc159:*ppi2* samples that were not present in NTAP:WT. For NTAP:WT only low molecular mass signals around 20kDa were observed and corresponded to monomeric NTAP.

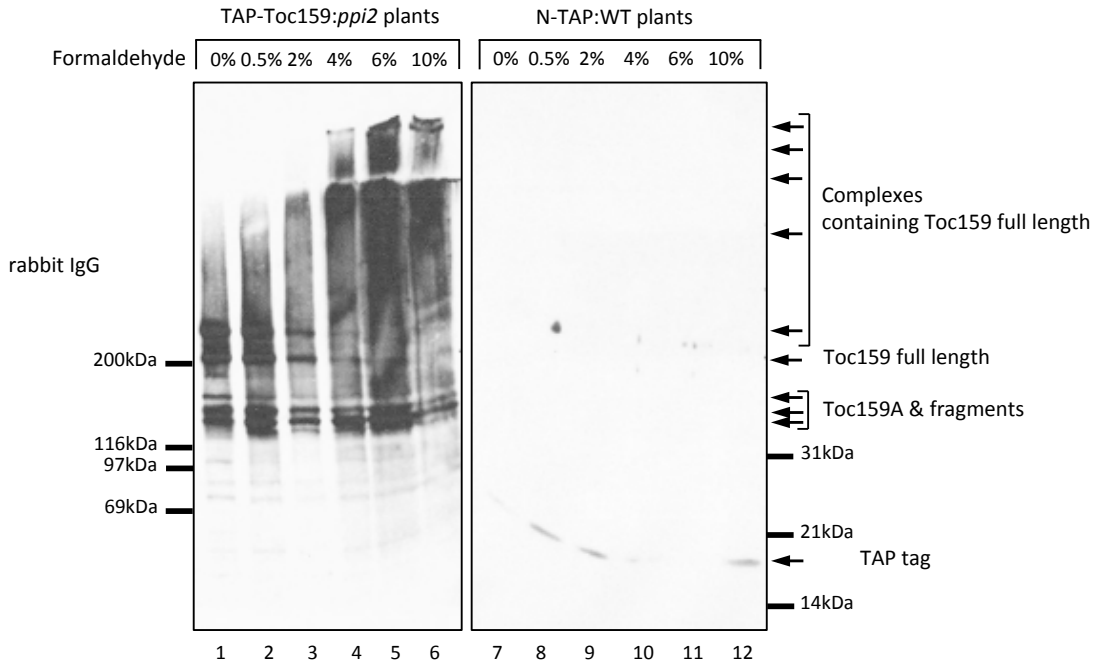


Figure 12. Optimization of formaldehyde concentration for use as a protein crosslinker.
Three week-old TAP-Toc159:ppi2 and N-TAP:WT whole Arabidopsis plants were immersed in a range of formaldehyde solutions as indicated (0-10%). After incubation, formaldehyde treated plants were washed and remaining formaldehyde quenched with glycine. Total protein was extracted and dissolved in SDS sample buffer and separated on a 4-16% SDS-PAGE gel and tranferred to nitrocellulose. Western blotting with rabbit IgG was used to detect TAP-Toc159 (lanes 1-6) and N-TAP (lanes 7-12).

In the absence of formaldehyde (Fig. 12, lane 1) and presence of 0.5% formaldehyde, 2 major bands were observed above 200 kDa and additional bands were observed in the 150kDa region. The band just above 200 kDa most likely corresponds to monomeric, full length Toc159 known to migrate at this relative mass on SDS-PAGE gels. The band migrating above it presumably represents a Toc159 complex. The

bands migrating around 150kDa most likely represent the separate, soluble TAP-Toc159A resulting from TAP-Toc159. There was an apparent gradual increase in the extent of crosslinking from 2% to 6% formaldehyde (lanes 3-5), which was visible from the shift of native TAP-Toc159 to higher molecular weight smears. Toc159 has been identified in a 880 – 1000 kDa Toc supercomplex (Chen & Li, 2006). Although there were no molecular mass markers available to indicate the mass range of the smears on this Western blot, they might contain the Toc supercomplex. Interestingly, the bands around 150 kDa remained relatively unchanged at the higher concentrations of formaldehyde. In addition, at 6 and 10% formaldehyde increasingly larger aggregates presumably were formed and remained inside the stacking gel (lanes 5 and 6).

In summary the use of formaldehyde in the range of 2% to 6% yielded crosslinked protein complexes of high molecular mass, indicating that the vacuum infiltration method worked efficiently to infiltrate formaldehyde crosslinker into the intact plants.

2.2.6 Crosslinking Toc159A-TAP complexes using formaldehyde

The previous experiment suggested that 4% formaldehyde might be a useful concentration to obtain crosslinked complexes without compromising protein recovery. Toc159A-TAP:WT plants and NTAP:WT plants were subjected to chemical crosslinking by 4% formaldehyde under vacuum infiltration. Total protein extracts were prepared from both crosslinked and non-crosslinked plants, separated on a 4-16% gradient BN-PAGE gel and transferred to nitrocellulose for Western blotting. The anti-CBP antibody was used to detect TAP tag specific complexes (Fig. 13). Anti-Toc159A was used to detect endogenous Toc159A and to confirm the results of the anti-CBP antibody.

Western blotting of BN-PAGE separated, crosslinked Toc159A-TAP:WT samples with anti-CBP revealed several higher molecular bands, marked by asterisks (Fig. 13, lane 2 & 5). Of these at least one was unique for crosslinked Toc159A-TAP:WT. However, there were also complexes detected in non-crosslinked Toc159A-TAP:WT samples (Fig. 13, lanes 1 & 4). No large molecular mass bands were detected using anti-CBP in the NTAP:WT sample that had been subjected to crosslinking (Fig. 13, lane 3).

The second Western blot using anti-Toc159A antibodies revealed additional bands identified along with those identified by anti-CBP antibodies. Some of these additional bands may correspond to endogenous Toc159 and Toc159A, respectively. This possibility is supported by the Western blotting of the NTAP:WT extract with anti-Toc159A antibody that revealed the bands that corresponded to endogenous full length Toc159 as well as Toc159A and their respective complexes (Fig. 13, lane 3).

Most likely, the lowest molecular mass bands in the NTAP:WT lane corresponds to Toc159A.

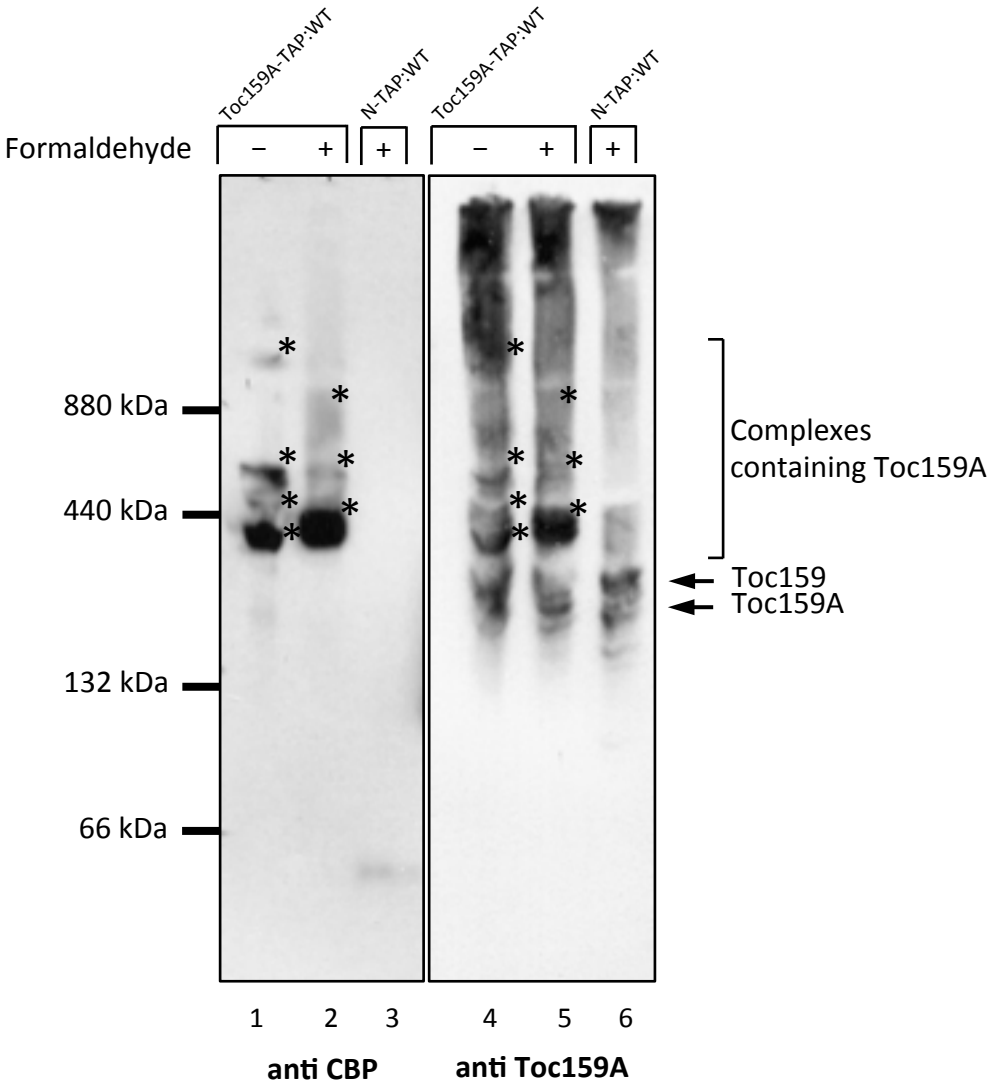


Figure 13. Identification of 4% formaldehyde crosslinked Toc159-TAP complex by BN-PAGE.

Three week-old Toc159A-TAP:WT and N-TAP:WT plants were immersed in 4% formaldehyde solution or PBS buffer and subjected to vacuum infiltration for 5 min. The formaldehyde solution was decanted and the plants washed immediately. Total protein extracts from Toc159A-TAP:WT plants incubated with PBS (1) and Toc159A-TAP:WT and N-TAP:WT plants treated with 4% formaldehyde (2 & 3) were separated on a 4-16% gradient Blue Native PAGE gel and transferred to nitrocellulose. The Western blot using a TAP tag specific antibody (anti CBP) identified several potential higher molecular mass complexes, indicated by the asterisks, both in non-crosslinked (1) and crosslinked (2) samples. The absence of these bands in negative control N-TAP:WT plants treated with crosslinker (3) indicated that Toc159A-TAP containing bands were specific. On the right hand panel, an identical Western blot was probed with anti Toc159A antibody, which revealed the endogenous Toc159A as well as transgenic Toc159A-TAP in the three samples (4, 5 & 6).

2.2.7. IgG affinity purification of formaldehyde crosslinked Toc159A-TAP complexes

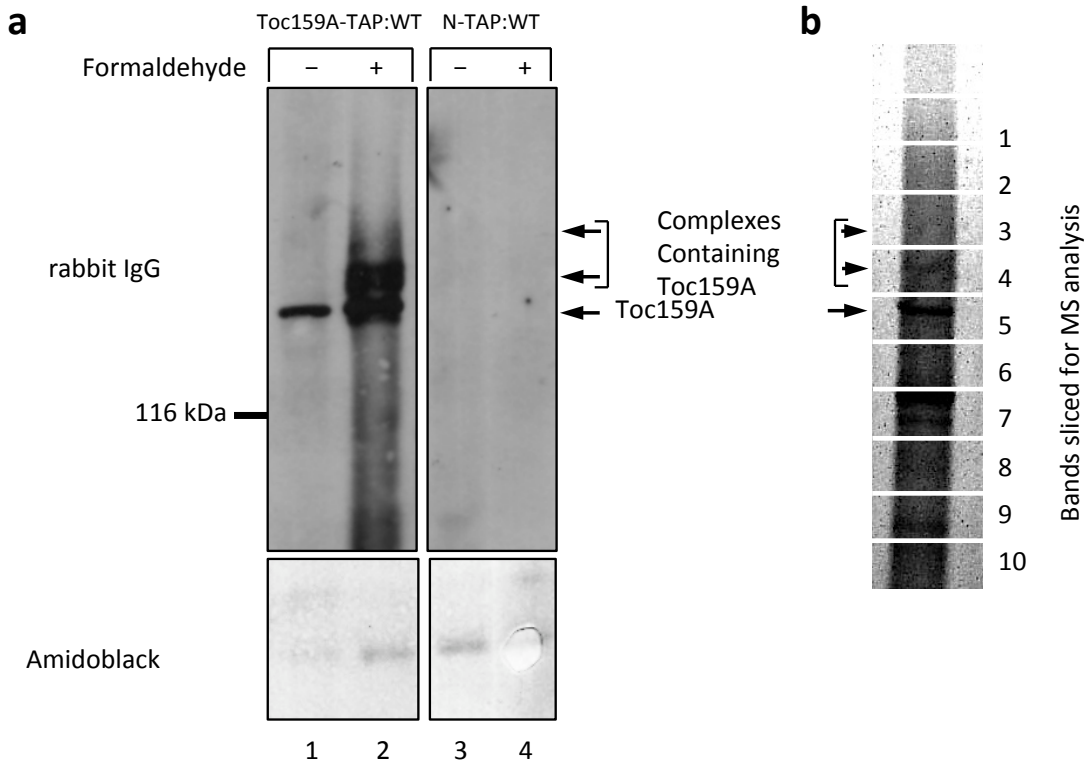


Figure 14. IgG affinity purification of formaldehyde crosslinked Toc159A-TAP complexes.
a) Toc159A-TAP:WT and control N-TAP:WT plants were incubated with a 4% formaldehyde solution or PBS buffer and subjected to vacuum infiltration for 5 minutes. Total protein was extracted using buffer containing 0.75% Triton X-100. The extract was clarified by centrifugation at 100,000xg. The supernatant was incubated with human IgG-sepharose and allowed to bind on a rotating shaker over night. The resin was washed and eluted with glycine buffer pH2.2. The eluates were separated on a 4-16% SDS-PAGE and transferred to nitrocellulose. Western blot was carried out using anti-rabbit IgG to identify Toc159A-TAP containing complexes. In the absence of 4% formaldehyde crosslinking (1) a single band corresponding to the Toc159A-TAP monomer was observed whereas an additional band of higher molecular mass was present in crosslinked sample (lane 2). b) Sypro-RUBY® staining of the Toc159A-TAP:WT crosslinked eluate. The eluate was separated on a 4-16% SDS gel and stained with Sypro-RUBY®. The lane of interest was then sliced in to eleven pieces of equal size for in gel trypsin digestion and extraction of tryptic- peptides for MALDI-TOF mass spectrometric analysis.

Toc159A-TAP:WT and NTAP:WT plants were subjected to 4% formaldehyde crosslinking as described in the previous paragraph 2.2.6. Total proteins were extracted from crosslinked as well as non-crosslinked plants (see. 4.2.4) and subjected to immunopurification as described (see. 4.2.9.3). Small aliquots of the

0.2M glycine (pH 3.0) eluates were separated on a 4-16% gradient SDS-PAGE gel and transferred to nitrocellulose. Western blotting with rabbit IgG was used for TAP tag detection (Fig. 14, panel a). Crosslinked Toc159A-TAP:WT eluate (lane 2) gave a major, slightly higher mass band in addition to Toc159A-TAP monomer which was also present in the non-crosslinked eluate (lane 1); both bands were absent from NTAP:WT eluates (lanes 3 & 4).

The remainder (90%) of the eluate was separated on an identical 4-16% gradient SDS-PAGE and stained with Sypro-RUBY® (Fig. 14, panel b). The gel was sliced into 11 pieces as shown and in-gel tryptic digestion was performed; peptides were eluted from the gel pieces and sent to the laboratory of Prof. Sacha Baginsky, at Martin-Luther Universität Halle -Wittenberg, where Dr. Birgit Agne performed the Mass spectrometry experiments and analysed the results (see. 4.2.9.4).

2.2.8. Results from Mass spectrometry analysis

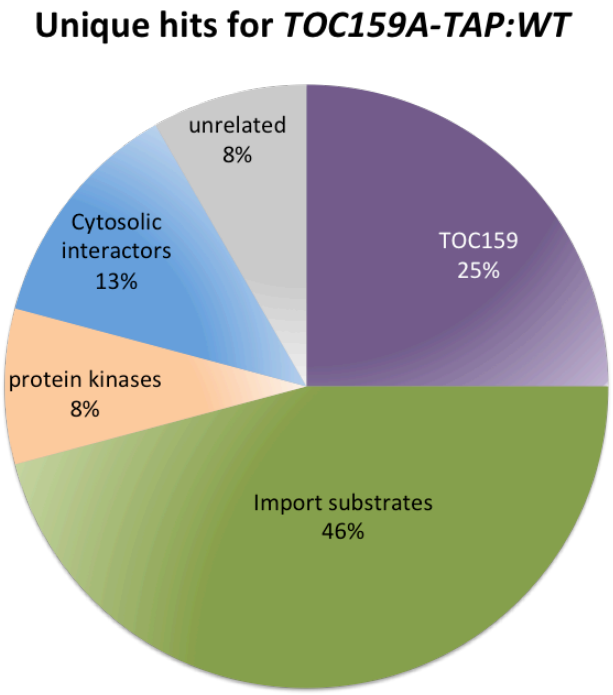
Eluates of the IgG affinity purification experiment of crosslinked Toc159A-TAP:WT were subjected to MALDI-TOF/MS analysis. The results of the mass spectrometric analyses of these eluates were compared with those of non-crosslinked *TOC159A-TAP:WT* as well as those of crosslinked *NTAP:WT*. Those hits that were unique for the *TOC159A-TAP:WT* eluate belonged either to Toc159 itself, were considered candidates for Toc159 interacting proteins or potential contaminations (Fig. 15). Based on their known or predicted molecular functions and cellular destination the candidate proteins were classified into chloroplast targeted potential Toc159 'import substrates', potential cytosolic interactors, chloroplast envelope and cytosolic protein kinases and other unrelated 'possible contaminants'.

The chloroplast targeted 'potential client' protein group includes proteins that are directly or indirectly involved in photosynthesis, chloroplast biosynthetic processes, plastid regulation and chloroplast protein synthesis. All of them are encoded by nuclear genome and synthesized in the cytosol as preproteins. We propose that their transit peptides might interact with Toc159A in the early stages of the import process.

The cytosolic interactors group contains the co-chaperone Hsp70 that is involved in cytosolic protein folding, a methyltransferase involved in translation and has ribosome binding properties and CAM5, a member of the calmodulin family.

Two kinases were identified that are new candidates for Toc159A phosphorylation.

There were two unrelated proteins, H+ ATPase and DOT2, which are possible contaminants because of their plasma membrane localization.



Accession	Description	Coverage	Peptides
AT4G02510	ATTOC159	5.72	6
AT5G04140	GLU1 _ glutamate synthase 1	1.36	2
AT2G41220	GLU2 _ glutamate synthase 2	0.92	1
AT5G17920	ATMS1 _ Cobalamin-independent a.a synthase	1.31	1
AT1G31330	PSAF _ photosystem I subunit F	4.98	1
AT1G52230	PSAH2 _ photosystem I subunit H2	7.59	1
AT3G50820	PSBO2 _ photosystem II subunit O-2	3.63	1
AT3G48870	HSP93-III _ Clp ATPase	1.41	1
AT3G01370	ATCFM2 _ CRM family member 2	1.48	1
AT4G10340	LHCB5 _ light harvesting complex	5.71	1
AT3G27850	RPL12-C _ ribosomal protein L12-C	7.49	1
AT5G03320	Unknown protein kinase superfamily protein	3.1	1
AT4G26540	LRR receptor-like protein kinase family protein	1.01	1
AT5G02500	AT-Hsp70 cytosolic cochaperone	3.07	1
AT1G64600	Unknown methyltransferase; copper ion bindin	1.86	1
AT2G27030	CAM5 _ calmodulin 5	14.16	1
AT1G80660	H(+)-ATPase 9	1.59	1
AT5G16780	DOT2, MDF _ SART-1 family	1.1	1

Figure 15. Mass spectrometry results.
The hits from crosslinked/non-crosslinked *TOC159A-TAP:WT* affinity purification were compared with those of crosslinked *NTAP:WT*, only the hits that were unique to *TOC159A-TAP:WT* are presented in the table. The proteins were classified according to gene ontology functional annotation on TAIR (The Arabidospsis Information Resource, www.arabidopsis.org) mauve: Toc159; green: chloroplast targeted proteins; orange: protein kinase family proteins; blue: cytosolic interactors; grey: unrelated proteins.

2.3. *In vitro* experiments with Toc159A

2.3.1. Extraction and purification of stable recombinant Toc159A

The acidic domain of Toc159 has been characterized *in vitro* and shown to belong to the group of intrinsically disordered proteins (Richardson et al., 2009). In earlier work from this laboratory recombinant Toc159A with a C-terminal hexahistidinyl tag (Toc159AHis-6x, $Toc159_{1\text{-}740}$-His-6x)) was used for experimentation. However, it was massively degraded during and after purification (Fig. 16) complicating *in vitro* experiments.

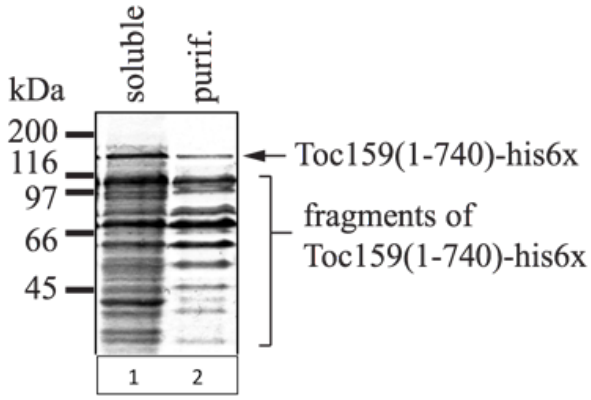


Figure 16. Instability of recombinant C-terminally His_6-tagged Toc159A (Toc159A-His6x or Toc159(1-740)-his6x).
The recombinant Toc159A-His6x was expressed as a soluble protein in *E.coli* BL21(DE3) cells (left lane) and purified by Ni-NTA column chromatography (right lane) . Note the presence of a large number of Toc159A-His6x fragments in the purified fraction. (This coomassie stained gel figure was published in Agne et al., 2010).

However, it had been demonstrated that an amino terminal His_6 tag largely stabilized the recombinant protein (Richardson et al., 2009). The protocol to purify the recombinant 6xHis-Toc159A was adapted from Richardson et al. (2009) as described in the section 4.2.3.1 The recombinant protein was expressed in *E. coli* BL21 (DE3) cells using the pET21b-6xHis-Toc159A plasmid, lysed using a French press and subjected to Ni-NTA affinity chromatography on a ÄKTAPrime® FPLC system. The recovery of recombinant 6xHis-Toc159A was followed throughout the procedure by taking small aliquots from each step and carrying out SDS-PAGE analysis and anti-Toc159A Western blotting (Fig. 17, panels a and b).

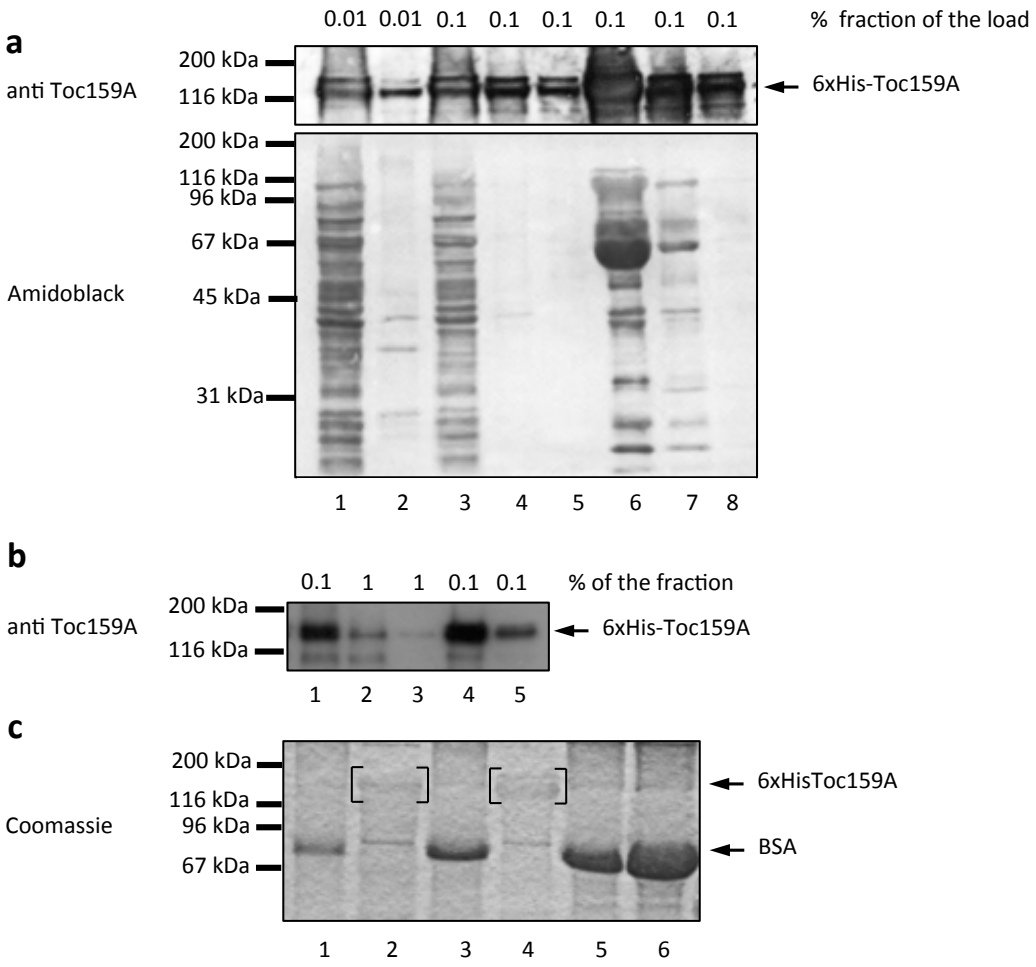


Figure 17. Production and purification of recombinant Toc159A with a N-terminal hexahistidyl-tag (6xhis-Toc159A)

a) The BL21(DE3) *E.coli* cells expressing 6xhis-Toc159A were cultured in a large volume. The cells were lysed using a French press. The crude extract was centrifuged. The clear supernatant was used for purification of soluble 6xhis-Toc159A. It was applied to a Ni-NTA column in an ÄKTAPrime® FPLC system. The clear extract (lane 1) pellet (lane 2), flowthrough (3), first (4) and last (5) wash fractions, along with three Ni-NTA eluates (6-8) were analyzed by SDS-PAGE followed by Western blotting. The Western blot using Toc159A antibodies indicated that the eluates contained a large proportion of the soluble recombinant 6xhis-Toc159A. However, Amido black coloration indicates impurities and degradation products. b) The Ni-NTA eluates were subjected to a second step of purification using DEAE anion exchange chromatography taking advantage of the strongly acidic isoelectric point of Toc159A. Lane 1: Ni-NTA eluates pooled (input for DEAE column), lane 2: first wash, lane 3: last wash, lane 4: DEAE eluate 1, lane 5: DEAE eluate 2. c) The eluates of the DEAE column were concentrated using a 35kDa cut-off molecular mass filter. The concentrated protein was analyzed by SDS-PAGE gel alongside 1, 5, 10 and 20µg of bovine serum albumin (lanes 1, 3, 5 & 6 respectively). Note that Toc159A is highly acidic and therefore poorly stained by Coomassie blue. Lane 2 and 4 contains 1 and 2 µL of the concentrated 6xHis-Toc159A.

The eluates of Ni-NTA affinity chromatograph step contained more than 10% of the total Toc159A from the bacterial extract as estimated from Fig. 17 (compare lanes 1 and 6). However, the eluates (6-8) still contained lower molecular weight impurities and degradation products that were visible either by Amidoblack staining or Western blotting. Therefore, a second purification step using a DEAE-ion exchange column was carried out. The eluates of the DEAE column were pooled and concentrated using an Amicon ultrafiltration device with a cut-off of 35kDa. Aliquots of the concentrated proteins were run on a SDS-PAGE along with known amount of BSA and the gel was stained with Coomassie blue (Fig. 17b).

2.3.2. Effect of Toc159A on chloroplast import of the precursor of the small subunit of RuBisCO

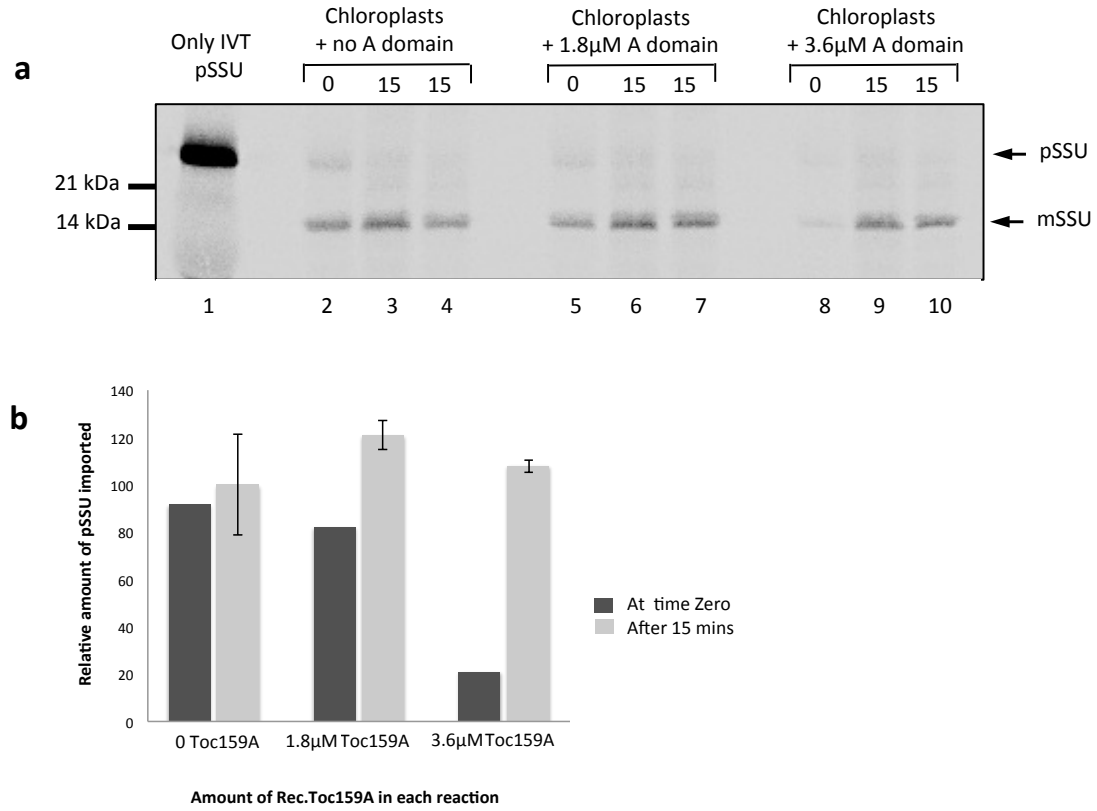


Figure 18. Effect of Toc159A on *in vitro* import of pSSU.
a) *In vitro* chloroplast protein import reactions were carried out using synthetic 35S-labeled Rubisco small subunit preprotein as a substrate. Purified, recombinant 6xHis-Toc159A was added in increasing concentrations as indicated. The import reactions were stopped at time zero (lanes 2, 5 & 8) or after 15 minutes (lanes 3, 4, 6, 7, 9 & 10). The chloroplast proteins were analyzed by SDS-PAGE followed by Coomassie blue staining. The dried gel was exposed to a Phosphorimager plate. b) The Phosphorimager quantification of the import reactions was plotted in a bar chart. The import at 0 minutes is shown as a dark grey bar. The light grey bars with standard deviation indicate two independent import reactions after 15 minutes.

The acidic domains of the members of the Toc159 GTPase family have been shown to bind certain classes of preproteins selectively (Inoue et al., 2010). To test whether binding to Toc159A influences chloroplast import of preproteins, import reactions into isolated Arabidopsis chloroplasts using ^{35}S-labeled *in vitro* synthesized pSSU as the substrate were carried out in the presence of 0, 1.8 and 3.6 µM of recombinant 6xHis-Toc159A. The reactions were stopped after 0 and 15 minutes (Fig. 18). After the import reactions the chloroplasts were separated by SDS-PAGE. The gels were colored with Coomassie blue, dried and exposed to a Phosphorimager plate. The Phosphorimager data indicated that ^{35}S-Met labeled pSSU had been imported and converted to lower molecular mass SSU (Fig. 18a).

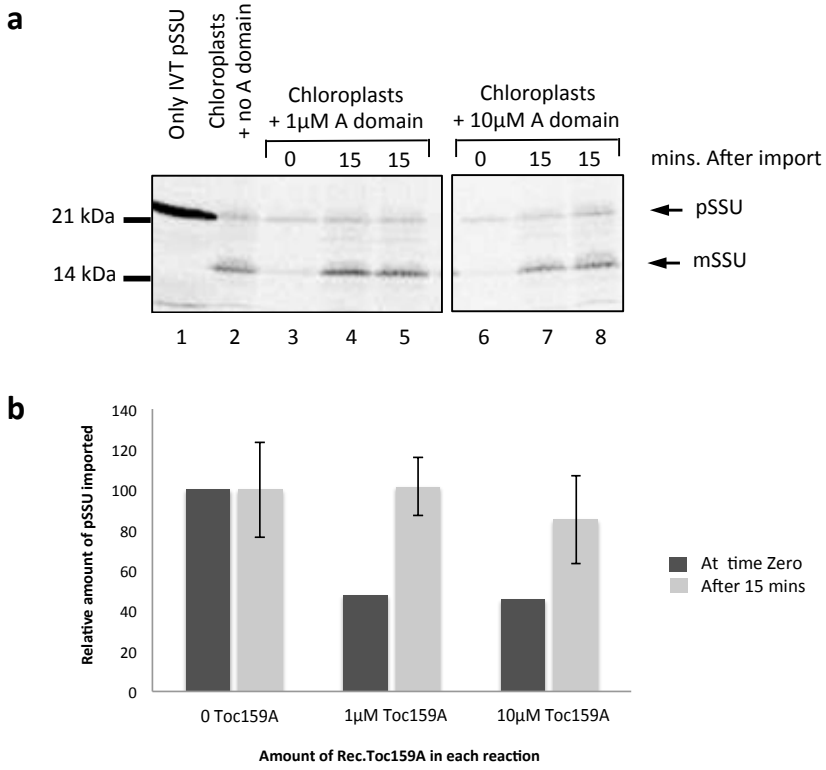


Figure 19. Effect of Rec. Toc159A in *in vitro* import of pSSU.
a) *In vitro* chloroplast protein import reactions were carried out using synthetic 35S-labeled Rubisco small subunit preprotein as a substrate. 0 (lanes 1,2 & 3), 1 (lanes 6, 7 & 8) or 10µM (lanes 6, 7 & 8) recombinant 6xHis-Toc159A was added. The import reactions were stopped at the end of 15 minutes. The chloroplasts were reisolated and separated by SDS-PAGE. The gel was dried and exposed to a Phosphorimage plate. b) The Phosphorimager quantification of the import reactions was plotted in a bar chart. The import at 0 minutes is shown as a dark grey bar. The light grey bars with standard deviation indicate two independent import reactions after 15 minutes..

The bands were quantified, and standard deviations calculated where possible; the results were plotted on a bar graph (Fig. 18b). The average of mature SSU bands quantified from import after 15 mins with no recombinant Toc159A was taken as 100 per cent, quantifications from rest of the lanes were represented in comparison to that.

The amount of imported mature SSU was not significantly changed by the absence or presence of Toc159A after 15 minutes of import. However, mature SSU detected after 0 min of import was strongly reduced by increasing concentrations of 6xHis-Toc159A.

Another set of import reactions was carried out using a higher concentration of 6xHis-Toc159A (Fig. 19a). This time 0, 1µM (equivalent to 3.3µg) and 10µM (equivalent to 33µg) 6xHis-Toc159A were added to the pSSU import reactions. The import reactions were stopped after 0 and 15 minutes. After the import reactions the chloroplasts were analyzed by SDS-PAGE. The gel was colored with Coomassie blue, dried and exposed to a Phosphorimager plate. The amount of imported mature SSU was not significantly changed by the higher concentration (10µM) of 6xHis-Toc159A after 15 minutes of import. However, mature SSU detected after 0 min of import again was strongly reduced by the presence of 6xHis-Toc159A. (Fig. 19b); it appears that 6xHisToc159A had no effect on pSSU import at this time point.

2.3.3. Phosphorylation of Toc159A in a cell free reticulocyte lysate

2.3.3.1. Recombinant Toc159A is phosphorylated in in vitro translation mix

During earlier studies in our laboratory Agne, B et al., it was discovered that Toc159A is hyperphosphorylated. However, the purpose of the phosphorylation remained mysterious. One of our hypotheses was that phosphorylation might be implicated in the regulation of preprotein import by selectively enhancing the translation of particular subsets of mRNAs in the cytosol.

We set up an experiment to analyze the effect of both phosphorylated and non-phosphorylated recombinant 6xHis-Toc159A on *in vitro* translation reactions of chloroplast targeted pSSU. Recombinant 6xHis-Toc159A was phosphorylated by Casein Kinase 2 *in vitro*. However, neither phosphorylated nor the non-phosphorylated 6xHis-Toc159A had an effect on *in vitro* translation (IVT) efficiency of pSSU (data not shown). Because no difference observed, we tested whether 6xHis-Toc159A was phosphorylated in the reticulocyte lysate during IVT of pSSU.

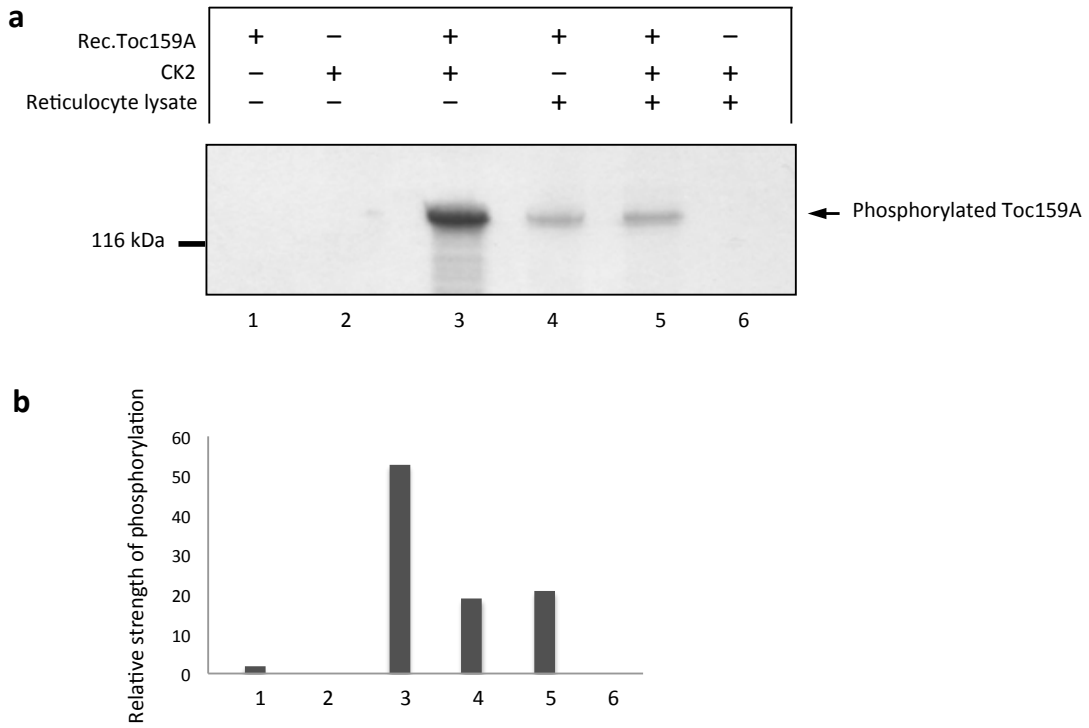


Figure 20. Phosphorylation of Toc159A in cell free reticulocyte lysate.
a) One microgram of the recombinant His_6-Toc159A was added to *in vitro* phosphorylation reactions containing CK2 (3), cell free reticulocyte lysate (4) or both CK2 and reticulocyte lysate (5) in the presence of 33P γ-ATP . The phosphorylation experiments were analyzed by SDS-PAGE followed by Coomassie staining and Phosphorimager quantification. b) Phosphorylation 6xhis-Toc159A under the different conditions were quantified using a Phosphorimager and plotted in a bar graph.

To determine the phosphorylation status of Toc159A in reticulocyte lysate, ^{33}P γ-ATP was added along with recombinant 6xHis-Toc159A. The reactions were separated on a SDS-PAGE gel. The gel was stained with Coomassie blue, dried and subjected to Phosphorimager analysis. Clearly, 6xHis-Toc159A was phosphorylated by the reticulocyte lysate (Fig. 20, lane 4). To compare, Toc159A was incubated with CK2 in the presence of ^{33}P γ-ATP (Fig. 20, lane 3) was used to compare the extent of phosphorylation. Phosphorimager quantification showed that phosphorylation was around 3-times as high with CK2 when compared to reticulocyte lysate.

2.3.3.2. Comparing the recombinant wild type and non-phosphorylatable 6xHis-Toc159A

In order to test the effects of Toc159A phosphorylation in different experimental settings, we engineered a plasmid construct to bacterially overexpress non-phosphorylatable TOC159A (termed Toc159A w/o ST;without serine or threonine), in

which all the serine and threonine (total of 67 sites) were replaced by alanine. The coding sequence for the Toc159A w/o ST was amplified using primers for Gibson cloning from a full length synthetic Toc159 cDNA (synthesized by Genecust, Luxembourg (see 4.1.7 & 4.2.1) in which all serine and threonine codons had been replaced by alanine codons. The recombinant 6xHis-Toc159A w/o ST protein was expressed and purified using the same procedures that were used for 6xHis-Toc159A. Prior to use of the purified Toc159A w/o ST in experiments it was important to A) check that mutation of serine and threonine to alanine did in fact abolish Toc159A phosphorylation and B) thatToc159A w/o ST was still recognized by anti-Toc159A antibodies in Western blotting experiments.

Both recombinant 6xHis-Toc159A and 6xHis-Toc159A w/o ST was incubated *in vitro* with commercial CK2 in the presence of 33P γ-ATP. The reactions were separated by SDS-PAGE followed by Coomassie blue staining followed by exposition to a Phosphorimager plate. The Phosphorimage shows that wild type recombinant 6xHis-Toc159A was phosphorylated by CK2, while the recombinant Toc159A w/o ST was not (Fig. 21). Western blotting of the same reactions, showed anti-Toc159A antibodies recognize both 6xHis-Toc159A and 6xHis-Toc159A w/o ST.

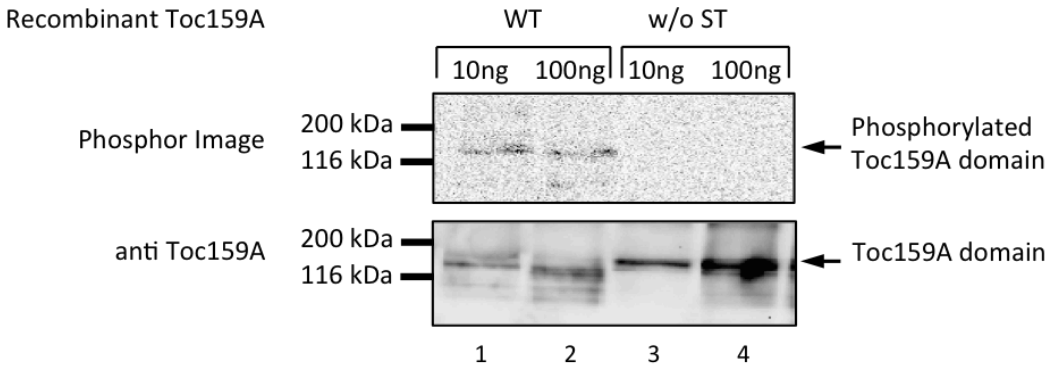


Figure 21. *In vitro* phosphorylation test of non-phosphorylatable 6xHis-Toc159A lacking serine and threonine residues that were substituted by alanine (6xHis-Toc159A w/o ST)
10 or 100 ng of wild type 6xHis-Toc159A or synthetic 6xhis-Toc159A w/o ST were subjected to phosphorylation by CK2 in the presence of 33P γ-ATP . Proteins in the reactions were separated by SDS-PAGE, stained with Coomassie blue and analyzed using a Phosphorimager. The Phosphorimager picture (top panel) shows phosphorylation only of the wild type 6xhis-Toc159A (lanes 1 and 2) but not of 6xHis-Toc159A w/o ST (lanes 3 and 4); Western blotting using the anti-Toc159A serum (bottom panel) confirms the presence of recombinant protein in all four reactions.

2.3.3.3 Effect of Toc159A phosphorylation on in vitro preprotein translation

To test the effect of Toc159A phosphorylation on preprotein *in vitro* translation, we chose three chloroplast targeted nuclear synthesized proteins. These three proteins belong to three developmental categories according to Teng et al., (Teng et al., 2012). This study demonstrated *in vitro* that proteins such as RuBisCO small subunit (pSSU) were favored for import by chloroplast from leaves at an early developmental stage and proteins such as pTic40 were preferred by chloroplast from leaves at a late developmental stage, whereas proteins such as pE1α did not show any age dependent preference.

It is well known that Toc159 is involved in chloroplast biogenesis. This suggests that Toc159 is mainly involved in the import of substrates required at early developmental stages (Bischof et al., 2011; Inoue et al., 2010). To test the effects of phosphorylatable and non-phosphorylatable Toc159A on preprotein translation, 0, 1 and 10µM of purified, recombinant 6xHis-Toc159A and 6xHis-Toc159A w/o ST, respectively, were added to cell free translation reactions of pSSU, pTic40 and pE1α in reticulocyte lysate in the presence of ^{35}S-methionine.

The *in vitro* translation reactions were separated by SDS-PAGE. The gel was stained with Coomassie blue, dried and analyzed using a Phosphorimager. To normalize gel loading a protein band contained in the reticulocyte lysate was chosen. The ratio of signal intensity of the reference band to the Phosphorimager signal of the corresponding radioactive preprotein in the absence of any recombinant Toc159A was arbitrarily set to 100%. Experiments in the presence of recombinant Toc159A were expressed as percentage of the reference. The normalized values were plotted in a bar graph (Fig. 22 D, E, F).

The results of Phosphorimager analysis (Fig. 22) suggest that regardless of the nature of the preprotein, increasing amounts of non-phosphorylatable Toc159A w/o ST reduced the *in vitro* translation of preproteins.

Phosphorylatable Toc159A seemed to specifically enhance the *in vitro* synthesis of pSSU (Fig. 22, A and D) with increasing concentrations of Toc159A in the reaction while it had a little negative influence on pE1α synthesis (B & E) and even reduced the synthesis of pTic40 to 5-6 fold (C and F).

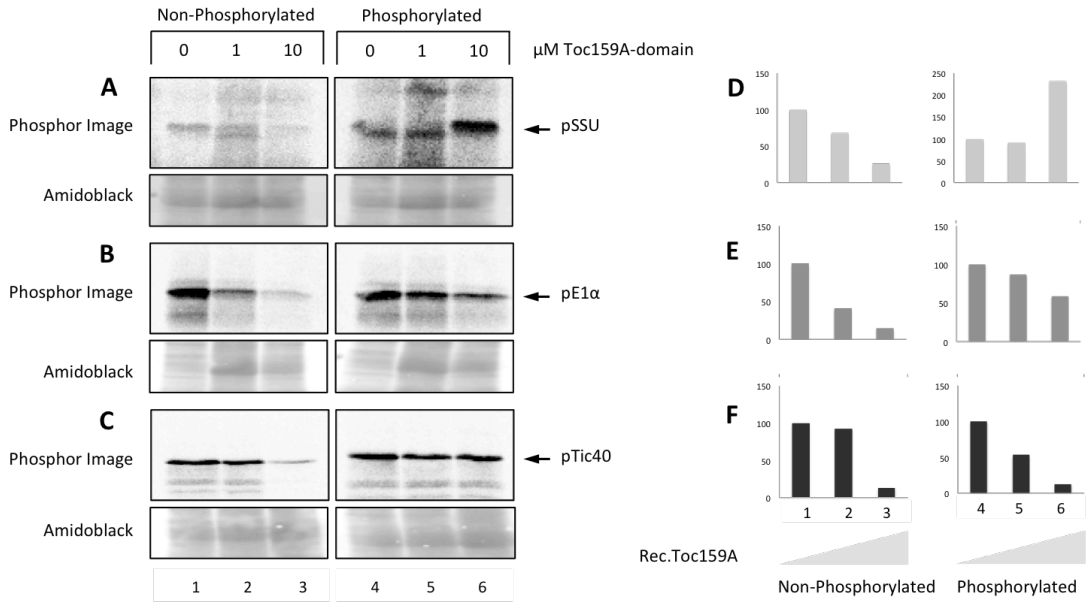


Figure 22. Influence of phosphorylated Rec.Toc159A in *in vitro* preprotein translation.
In vitro translation reactions of pSSU (Panel A), pE1α (Panel B) and pTic40 (Panel C) were carried out in cell free reticulocyte lysate in the presence of 35S-methionine. Purified 6xHis-Toc159Aw/o ST (lanes 1,2 and 3) and 6xHis-Toc159A (lanes 4,5 and 6) were added at 1 µM (lanes 2 and 4) or 10µM (lanes 3 and 6) to the *in vitro* reactions. Five percent of each of the translation reactions were analyzed by SDS-PAGE followed by Coomassie blue staining and Phosphorimager analysis. The phosphorimage of the IVT products were quantified and adjusted for the loading differences before plotting into bar chart. (Panels D, E and F)

3. Discussion

Toc159 was originally discovered as 86kDa protein in pea. Only much later was it discovered that the 86 kDa was only part of a much larger protein of 159 kDa) (Schnell et al., 1994; Waegemann & Soil, 1991). In addition to the GTP-binding (G-domain, Toc159G) and membrane insertion domain (M-domain, Toc159M), the 159 kDa protein had a large acidic N-terminal domain (A-domain, Toc159A). Ever since it's belated discovery the A-domain has fascinated researchers in the field yet it has not yielded all of its secrets. My supervisor and his collaborators continued to pursue this subject. In a recent publication (Agne et al., 2010) they demonstrated that Toc159A, is released from the full length Toc159. This paper suggested that separated Toc159A is present both as a soluble protein in the cytosol and as a membrane-bound protein at the chloroplast envelope. Furthermore, it was demonstrated that Toc159A exists a hyperphosphorylated protein both in the cytosol and at the chloroplast. Amongst other kinases, it is a very efficient substrate of casein kinase II (CK2). However, we currently do know neither the function of the A-domain neither the significance of the release of the A-domain nor that of its phosphorylation. During my thesis, I focused on the characterization of Toc159A complexes with the ultimate objective of identifying interacting proteins. In a parallel approach, I addressed two hypotheses regarding the function of Toc159A: A) that it may interfere with preprotein import into isolated chloroplasts and B) that it affects the synthesis of cytosolic preproteins prior to import into the chloroplasts.

First the question of Toc159A localization was addressed. While the study by Agne et al. (Agne et al., 2010)had shown that endogenous, released Toc159A was present in the soluble as well as chloroplast membrane fractions, it could not be excluded that it had been released from chloroplasts during the fractionation procedure. When a Toc159A YFP fusion protein (Toc159A-YFP) was expressed either transiently in isolated protoplasts or stably in transgenic plants YFP fluorescence was present in the cytosolic spaces between chloroplasts as well as in tight rings surrounding chloroplasts but not in the chloroplast stroma. The cytosolic localization is not surprising because the A-domain does not have any known targeting information and is not required for targeting of Toc159GM (lacking the A-domain) to the outer chloroplast membrane. The ring-like fluorescence is typical for chloroplast envelope localization. However, outer and inner membrane localization cannot be resolved using confocal laser microscopy. In the case of Toc159A-YFP, outer chloroplast membrane localization appears most likely. Given that Toc159A does not contain any known targeting information, the question of how Toc159A associates with chloroplasts may be asked. Toc159A is highly hydrophilic and lacks any predicted transmembrane domains. It therefore appears extremely probable that Toc159A extrinsically associates with the outer chloroplast membrane.

To shed further light on the nature of the chloroplast association of Toc159A thermolysin digestion was carried out on isolated protoplasts and chloroplasts of plants expressing Toc159A-TAP. Surprisingly, Toc159A-TAP in isolated chloroplasts was highly resistant to the protease that in the same experiment completely degraded Toc33, a preprotein import receptor at the chloroplast surface known to be sensitive to thermolysin. Moreover, purified recombinant 6xhis-Toc159A was completely degraded by thermolysin in solution. The thermolysin resistance of Toc159A-TAP in isolated chloroplasts may be explained in more than one way: the A-domain may have gained access to the intermembrane space between outer and inner chloroplast membranes which we would not be able to distinguish using confocal laser microscopy of Toc159A-YFP. While it cannot be completely ruled out, this scenario appears unlikely because of the lack of known targeting information in Toc159A that would probably be required for its entry into the intermembrane space. In an alternative scenario, Toc159A may be located extrinsically at the chloroplast surface in tight association with other proteins that would protect it against the protease activity. Hypothetical candidates for Toc159A interacting proteins at the outer membrane may be components of the Toc-complex or other currently unknown proteins at the outer chloroplast membrane. In yet another scenario, posttranslational modification, in this case hyperphosphorylation, may decrease the sensitivity of Toc159A to thermolysin or make it completely insensitive.

In this thesis I particularly addressed the hypothesis that Toc159A participates in protein-protein interactions. The A-domain was shown to behave as an intrinsically disordered protein and such proteins are known to participate in interactions with other proteins. I also hypothesize that phosphorylation of Toc159A may affect such interactions. Toc159A is vaguely related to the acidic ribosomal phosphoproteins P1/P2, so that we considered the possibility that Toc159A influences protein translation.

We used a variety of methods to address Toc159A protein complex formation including Blue Native PAGE (BN-PAGE), TAP-tag pull down, chemical crosslinking using dithiobis-succinimidyl propionate (DSP) and formaldehyde. BN-PAGE was carried out in view of identifying and isolating stable Toc159A complexes either of cytosolic or chloroplast envelope origin. While BN-PAGE clearly showed the presence of two bona-fide Toc159A complexes in a digitonin-solubilized total plant extract (Fig. 8, panels a and b), it was impossible to obtain sufficient material to carry out mass spectrometric protein identification. However, the stability of the two complexes by BN-PAGE suggested that affinity purification by TAP-tag purification should be feasible.

For this purpose, Toc159A-TAP:WT plants were separated into soluble, cytosol containing and chloroplast membrane containing fractions, respectively. The soluble

and detergent-solubilized chloroplast membrane fractions were subjected to human IgG-affinity chromatography (Fig. 9). Unfortunately, this procedure despite my best efforts did not yield the expected results. While the binding of Toc159A-TAP to the column appeared efficient, I was unable to detect Toc159A-TAP in the eluates specifically. This could be attributed either to degradation during the washing procedure or to inefficient elution.

At this point, I decided to include chemical crosslinking to stabilize Toc159A complexes and possibly increase the yield of interacting proteins. A series of different approaches were tested using either DSP or formaldehyde as crosslinking reagents. Both of the reagents are membrane permeable and can therefore be used on isolated protoplasts which would lead to crosslinking of Toc159A present either in the cytosol or at the chloroplast envelope membrane.

First, DSP was applied to isolated Toc159A-TAP:WT protoplasts that were subsequently separated into soluble, cytosol-containing and chloroplast membrane fractions and analyzed by SDS-PAGE followed by Western blotting using rabbit IgG to detect Toc159A-TAP (Fig. 10). While no bands that may correspond to a Toc159A-TAP complex were identified after DSP crosslinking of the soluble fraction, Toc159A-TAP in the chloroplast membrane fraction appeared to be shifted to a high molecular mass smear that failed to enter the running gel. The smear may represent a large molecular mass complex. However, a similar smear was also observed before and after DSP-crosslinking of wild type samples suggesting that the smears corresponded to non-specific aggregates.

Because the identification of an apparent DSP-crosslinked complex in Toc159A-TAP:wt protoplasts was not successful, vacuum infiltration of whole plants prior to fractionation into soluble, cytosol-containing and chloroplast membrane fractions was attempted (Fig. 11). The soluble and detergent-solubilized membrane fractions were separated by SDS-PAGE followed Western blot analysis after DSP-crosslinking revealing the presence of monomeric Toc159A-TAP migrating at the expected relative mass of around 116 kDa and an additional band migrating above 200 kDa. The 200 kDA band may represent a crosslinked complex and be identical with the smaller of the two complexes that was observed by Blue native-PAGE. However, this conclusion would benefit from two controls namely the omission of the DSP crosslinker and the treatment of an aliquot DSP crosslinked sample with reducing sample buffer to undo the crosslinking prior to complex analysis. In both cases, the disappearance of the 200 kDa band would be expected in the case of a *bona fide* Toc159A-TAP complex.

Finally, formaldehyde was used for chemical crosslinking, an old reagent that has recently been used for crosslinking in animal cell cultures. To explore suitable

formaldehyde crosslinking we initially used complement full length TAP-Toc159:ppi2 plants and NTAP:WT plants as a control for vacuum infiltration using increasing concetnrations of formaldehyde (Fig. 12). Analysis by Western blot showed that TAP-Toc159 were shifted increasingly to higher molecular mass smears by higher concentrations of formaldehyde while monomeric Toc159 disappeared. The higher molecular mass smears may contain the 880 kDa and 1 MDa complexes that have been reported in the literature. Based on the results of this experiment, a concentration of 4% formaldehyde appeared suitable for further experimentation.

In the following Toc159A-TAP:WT and NTAP:WT control plants were vacuum infiltrated using a 4% formaldehyde solution (Fig. 13). Total protein extracts were analyzed by BN-PAGE gels and revealed Toc159A-TAP specific band. Interestingly, similar bands were observed with and without 4% formaldehyde crosslinking. This may be due to the fact that some Toc159A-TAP complexes are stable even in the absence of formaldehyde (see also Fig. 8, panel b). Indeed, some of the bands seen in both bands may in fact be identical even though the molecular masses do not appear to be slightly lower in the experiment in Fig. 8 A band migrating below the 440kDa marker most likely represents the Toc159A-TAP monomer (see also Fig. 8, panel a). The first band migrating slightly more slowly in Figure 8, lane 1 may represent a Toc159A complex. To obtain sufficient material we proceeded to isolation of the 4% formaldehyde crosslinked complex by human IgG chromatography (Fig 14, panel a). Analysis of a small aliquot of crosslinked material by SDS-PAGE followed by Western analysis revealed the abundant presence of a higher molecular mass band after crosslinking (Fig 14, lane 2). The rest of sample was separated on SDS-PAGE gradient gel and stained with Sypro-RUBY. The lane was sliced into 11 pieces that were subjected to in gel tryptic digest. Peptides were eluted and subjected to mass spectrometry analyses.

This analysis led to the identification of 10 chloroplast proteins including PsaF, PsaH2, PsbO2 and Lhcb5 which are required for photosynthesis. They are all synthesized as preproteins in the cytoplasm and imported into chloroplasts. Therefore these proteins are probable substrates of the Toc159 import receptor. Their identification in the complex suggest that they interact with the A-domain of Toc159 prior to outer membrane translocation. This hypothesis is in agreement with recent findings by Dutta and colleagues (Dutta et al., 2014).

The mass spectrometric analysis identified two protein kinases. One is an uncharacterized serine/threonine kinase predicted to be localized in chloroplasts. The other is a predicted plasma membrane/ cytosolic LRR-receptor like protein kinase. Previous experimental studies have identified 43 phosphorylation sites in Toc159A. Most of these can be attributed to CK2 phosphorylation but other kinases such as SnRK2 have also been identified (Agne et al., 2010; Durek et al., 2010; Wang et al.,

2013). The implication of still other kinases in Toc159A phosphorylation is a clear possibility. The two kinase family proteins identified in this study have the potential to be such kinases.

We also identified the cytosolic cochaperones Hsp-70 that is known to be involved in the folding of the nascent proteins in the cytoplasm. In pea, Hsp70 has been shown to be part of the "Guidance complex" that targets preproteins to the TOC complex (Flores-Pérez & Jarvis, 2013; Lee et al., 2013; 2014). The result suggests that Hsp70 may mediate the interaction of the preprotein-containing guidance complex with the A-domain of Toc159. Potentially, this constitutes an early step of outer membrane translocation.

Taking advantage of a new, formaldehyde-based crosslinking strategy and subsequent immunopurification, it was possible to identify potential Toc159A-interacting proteins by mass spectrometry. However, it is also clear that these results need to be confirmed by further, improved isolation experiments with increased yield to achieve a higher number of mass spectrometry hits. Furthermore, the implications of any of the potential Toc159A interacting proteins in chloroplast protein import need to be confirmed using independent methods.

In the second part of my thesis, the hypotheses with regard to the function of Toc159A in chloroplast preprotein import and protein translation were addressed. In earlier studies, recombinant C-terminally hexahistidinyl-tagged Toc159A was used. However, after purification from bacterial extracts this protein was strongly degraded. Therefore, we used N-terminally hexahistidinyl-tagged Toc159A (6xhis-Toc159A) that was purified in a two-step procedure employing Ni-NTA chromatography followed by DEAE-column chromatography. This resulted in preparations of Toc159A that were far more intact than previously. This was a prerequisite to use the recombinant protein in *in vitro* assays. Following the same protocol synthetic, non-phosphorylatable Toc159A in which all serine and threonine were replace with alanine (Toc159A w/o ST) was also purified successfully.

Recent publications suggest that preprotein selectivity of Toc159 and its homologs depends on the A-domains. (Dutta et al., 2014; Inoue et al., 2010) If this were the case, competitive inhibition of *in vitro* preprotein import of pSSU by the recombinant A-domain might be expected. However in the concentration range tested (up to 10µM) no effect was observed on the accumulation of the mature SSU at the 15 minute point. However, at the earliest time point (0') accumulation of mature SSU was reduced. While these results still need to be reproduced, they suggest an effect of the A-domain on the kinetics of pSSU import. It is well possible that the recombinant A-domain does

not faithfully recapitulate the function of the A-domain *in vivo*, and may therefore not fully unfold its activity.

The A-domain of Toc159 has a weak similarity with acidic ribosomal proteins, the pI values of which are comparable to that of Toc159A (pI 3- 4), are required for full translational activity and change ribosomal specificity towards different subsets of mRNAs (Bautista-Santos & Zinker, 2014; Tchórzewski, 2002) For this reason we hypothesized that the A-domain may affect translation of cytoplasmically synthesized chloroplast preproteins. We therefore tested the effect of the A-domain (both wild type and non-phosphorylatable Toc159A) on the *in vitro* translation of selected preproteins each corresponding to one of three distinct categories of preproteins: early stage (pSSU) late stage (pTic40) and age independent proteins (pE1α) that have been proposed by Teng et al., (Teng et al., 2012). I observed that wild type Toc159A enhanced "early stage" pSSU synthesis but decreased synthesis of "late stage" Tic40 and "age independent" pE1α. These data suggest that the A-domain could indeed affect mRNA specificity of the ribosome. Interestingly, this effect was only observed with wild type, phosphorylatable Toc159A. The non-phosphorylatable Toc159A w/o ST inhibited translation of all three preproteins. It is therefore imaginable that wild type Toc159A must be phosphorylated to a certain extent to enhance translation of preferred substrates. Our experiments showed that phosphorylation of wild type Toc159A occurred in the cell free reticulocyte lysate used for *in vitro* translation. Not surprisingly no phosphorylation of Toc159A w/o ST was observed in the reticulocyte lysate. However, we cannot exclude that the S and T to A substitutions changed the activity of Toc159A independently of phosphorylation. These are exciting preliminary results but much about Toc159A- dependent selectivity remains open to be discovered.

The findings of this study are summarized in a scheme (Fig.23); it also accounts for data from the literature whenever available.

According to the scheme Toc159A is released from full length Toc159 by an unknown protease. The separated Toc159A is implicated in cytosolic and chloroplast interactions. Based on the based on the in vitro translation, Toc159A in its phosphorylated state may stimulate the translation of the early stage pSSU whereas the translation of other types of chloroplast-targeted preproteins was inhibited. Such a mechanism may couple translation of early stage proteins with translocation into the chloroplast during chloroplast biogenesis.

The MS hits included cytosolic cognate heatshock protein Hsp70. Hsp70 is part of guidance complex that deliver preproteins to chloroplast (Flores-Pérez & Jarvis, 2013). However, since no other proteins of the guidance complex were

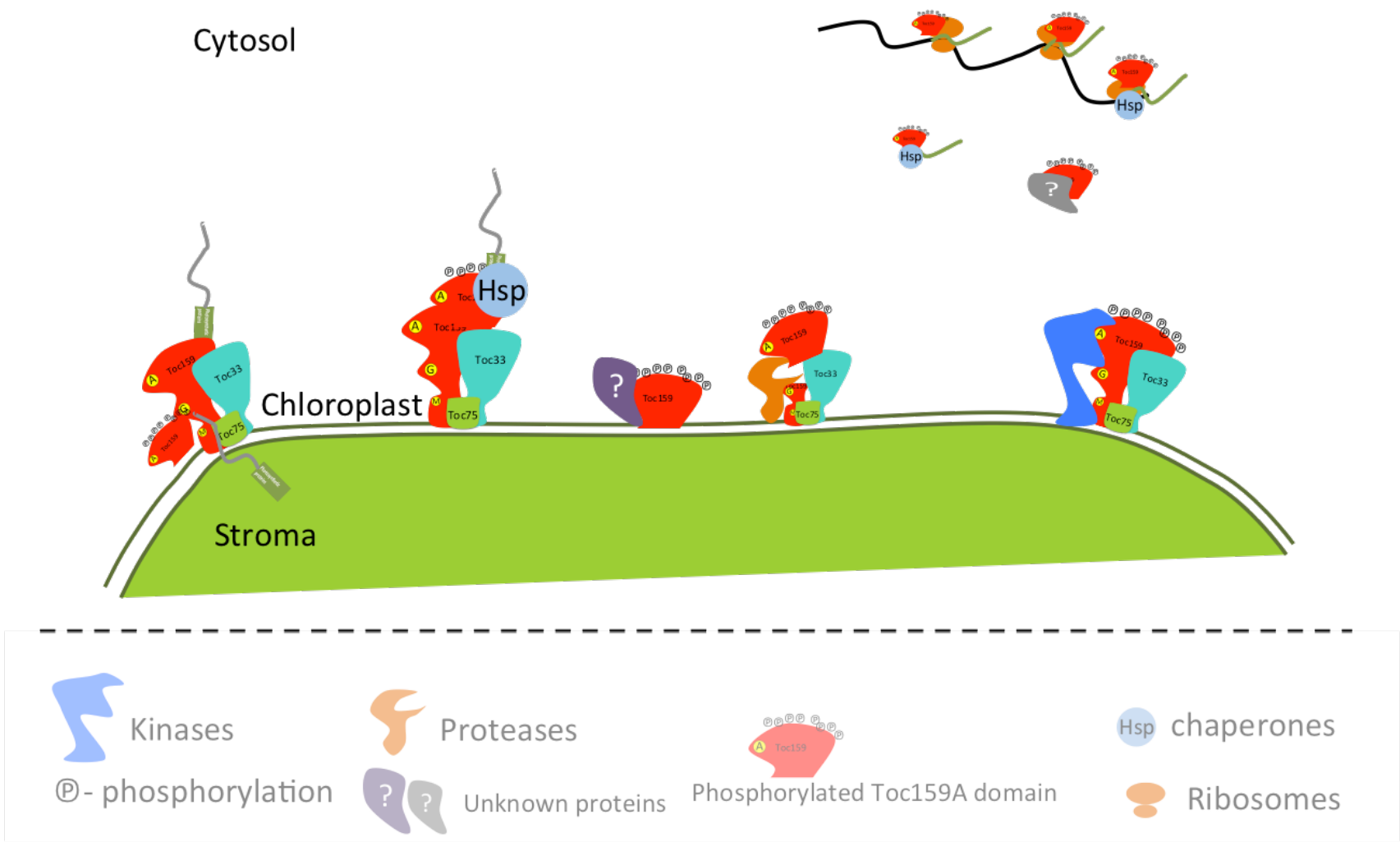


Figure.23. Schematic diagram of Toc159A functional interactions including that were identified in this study

found in the MS results, Hsp70 may represent the direct interaction partner of Toc159A whereas the other components were lost. Toc159A interaction with Hsp70 could strengthen the Hsp70 interaction with chloroplast surface, which might be critical for the rate of preprotein delivery.

It appears likely that the interaction of separated Toc159A with the chloroplast surface may involve interaction with the Toc complex. The thermolysin sensitivity data suggest that Toc33 is not a potential protease-protecting partner as it was digested by thermolysin. However, Toc75, a known interaction partner of Toc159, may protect the A-domain as it is known to be thermolysin insensitive. However, the MS data on Toc159A interacting proteins did not identify any components of the complex. It is therefore possible that unknown components of the chloroplast outer membrane protect separated Toc159A.

The MS data identified kinases as potential interaction partner of the separated Toc159A. It has been shown that the multiple kinases target the Toc159A. The kinases identified in this study may therefore represent exciting new research objects.

Future perspectives

The localisation of Toc159A domain to the chloroplast membrane has been confirmed by this study; however, future fluorescent co-localisation experiments could help understand the localisation in detail. This could be performed using fluorescent-tagged chloroplast outer and inner envelope proteins in the *TOC159A:YFP* plants. The cytosolic co-localisation can be analysed by co-localizing with non-targeted pSSU:GFP as well as fluorescent tagged cytosolic marker proteins. But all these experiments could only indicate the localisation, for identifying specific interactions we still need the confirmation from *in vivo* immuno pull down and mass spectrometry results.

We understand the MS analysis has to be repeated for more confident hits. At present the bottleneck in achieving this is mainly the low yield of pull-down eluates available for in-gel digestion and the exposure of samples to contaminants during multiple treatments. One way to overcome this issue could be to perform multiple pull-downs and pool the eluates to increase the recovered proteins before going for In-gel digestion. After finalizing the list of potential interacting proteins by mass-spectrometry the interaction with each candidate protein has to be tested for impaired function/functional interaction in single and double mutants and *in vitro* import reactions with single or double mutant chloroplasts.

In vitro translation with phosphorylatable and non-phosphorylatable Toc159A has shed light on the importance of Toc159A phosphorylation in preprotein interaction. To

understand the underlying mechanism, we can express the non-phosphorylatable Toc159 full-length protein in *ppi2* background and analyze the profile of proteins imported into chloroplasts in comparison to WT. If *TOC159w/oST:ppi2* were lethal, we can use chloroplasts from *TOC159GM:ppi2* plants in *in vitro* import reactions supplemented with phosphorylatable or non-phosphorylatable recombinant Toc159A.

4. Materials and methods

4.1 Materials

4.1.1. Antibodies

Anti-His_6 antibodies were purchased from Invitrogen, Life technologies, Thermo Fischer Scientific Inc., Waltham, MA, USA. Anti-CBP, from GenScript Inc., New Jersey, USA. Rabbit IgG antibodies from MP Biomedicals LLC, CA, USA. Human IgG used in TAP tag pull down assays was obtained from Sigma Aldrich, MO, USA.
Anti-Toc33 and anti-Toc159A have been described (Bauer et al., 2000; Hiltbrunner et al., 2001).

4.1.2. Arabidopsis lines

Wild type (wt) Arabidopsis plant always refers to *Arabidopsis thaliana* (L) var. wassilewskija. Toc159A-TAP Arabidopsis plant refers to wild type (wt) plants that were transformed with pCH7-Toc159 $_{1\text{-}764}$-TAP that was published earlier(Agne et al., 2010). Toc159A-YFP plants refer to wild type plants that were transformed with pEG101-3HA-$Toc159_{1\text{-}759}$-YFP construct.

4.1.3. Bacterial strains

DH5α, and One Shot ® OmniMAX ™ *E. coli* cells were purchased from Invitrogen, Life technologies, Thermo Fischer Scientific Inc., Waltham, MA, USA. BL21(DE3) *E. coli* cells were bought from New England Biolabs, Ipswich, MA, USA.

The *Agrobacterium tumefaciens* C 58 strain was used for Arabidopsis floral dip transformation.

4.1.4. Chemicals

Unless stated otherwise, the chemicals were purchased from Fluka chemicals GmbH.
^{35}S Methionine and γ ^{33}P- ATP were from Hartman Analytic GmbH, Braunschweig, Germany.

4.1.5 Equipment

The semi-automated hi-pressure chromatography system ÄKTAprime, hi-Trap Ni-NTA and hi-Trap DEAE columns were obtained from Amersham Biosciences, USA.

The steel ball mill used for grinding plant samples was supplied by Retsch, F.Kurt Retsch GmbH, Haan, Germany.
Large volumes of bacterial liquid culture was grown in a 5 lit fermenter supplied by INFORS AG, Switzerland.

4.1.6. Oligonucleotides

All the oligonucleotides were ordered from Microsynth AG, Balgach, Switzerland.
Synthetic full-length genes encoding wild type Toc159 or non-phosphorylatable Toc159 (all serine and threonine residues replaced by alanine) (see Fig. 23), were synthesized by GeneCust, Luxembourg.

4.1.7. Vectors

pET21 b was purchased from Novagen, Inc., Merck KGaA, Darmstadt, Germany.
The pET21b-His_6-Toc159A construct was kindly provided by Dr. Matthew D. Smith, Wilfrid Laurier University, Canada.
The gateway vectors pDONR221, pEarlyGate 101, pEarlyGate 102 were obtained from Invitrogen, life technologies, Thermo Fischer Scientific Inc., Waltham, MA, USA.

4.2. Methods

4.2.1. Plasmids used for protein expression in E.coli, *in vitro* transcription/translation, protoplast transient expression and stable plant transformation

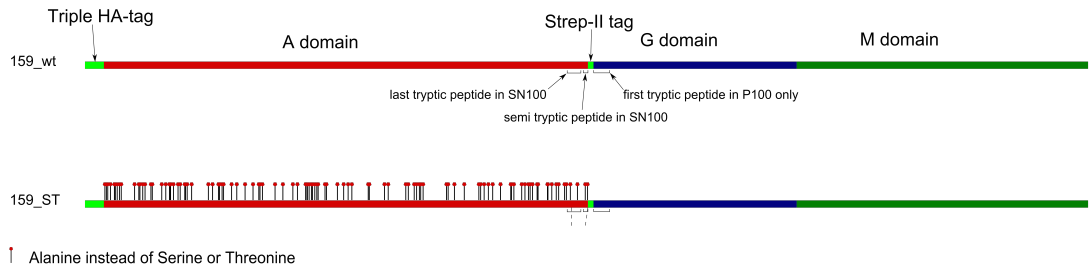


Figure.24. Synthetic Toc159 Primary structure
Toc159 wild type and serine threonine codons replaced for alanine were designed to study the function of phosphorylation on Toc159. For this study the sequences coding Toc159A acidic domain (1-759 a.a) were amplified from these constructs.

For fluorescent fusion protein construct pEG101 3HA-Toc159A-YFP (see. 4.2.1.5) and non-phosphorylatable recombinant protein construct pET21b His_6 –Toc159A w/o ST (see. 4.2.1.4) the region corresponding to acidic domain (1-759 a.a) were amplified from the 159_wt and 159_ST respectively along with triple HA tag for cloning purpose.

4.2.1.1 PCR amplification

PCR reactions were performed with *pfu* proof reading polymerase from Promega, the reaction conditions were set according to the manufacturers recommendation unless stated differently. Each reaction contained 200µM of each dNTP, 1.0µM upstream primer, 1.0µM downstream primer, DNA template varying from 10 – 200ng, *pfu* DNA Polymerase 1.25u/50µl and nuclease-free Water to a final volume of 50µl.

The following thermo cycler program was used and adapted to different primer pairs and templates:

Initial Denaturation	95°C	1–2 minutes	1 cycle
Denaturation	95°C	0.5–1 minute	
Annealing	42–72°C	30 seconds	25–35 cycles
Extension	74°C	2–4 minutes	
Final Extension	74°C	5 minutes	1 cycle
Soak	4°C	Indefinite	1 cycle

4.2.1.2. Gateway® cloning of 3HA-TOC159A$_{(1-759)}$ WT

BP reaction

The following reagents were added to an Eppendorf tube as recommended by the manufacturer. The attB-PCR product (=10 ng/µl; final amount ~15-150 ng) 1-7 µl, Donor vector pDONR 221 (150 ng/µl) 1 µl, TE buffer, pH 8.0 to 8 µl. To each reaction 2 µl of BP Clonase™ II enzyme mix was added, mixed well by vortexing, and briefly microfuged. The reaction mix was incubated at 25°C for 1 hour. Reactions were stopped by the addition of 1 µl of Proteinase K solution. The mixture was vortexed briefly and incubated at 37°C for 10 mins.

LR reaction

Reactions contained the Entry clone (50-150 ng) 1-7 µl, Destination vector either pEG101 or pEG102 (150 ng/µl) 1 µl, and TE buffer, pH 8.0 to 8 µl. After mixing the components 2µl of LR clonase was added to each reaction. The reaction was incubated at 25°C for 1 hour. Reactions were stopped by the addition of 1 µl of Proteinase K solution. The mixture was vortexed briefly and incubated at 37°C for 10 mins.

4.2.1.3. Gibson Assembly Cloning of pET21bHis$_6$ –TOC159$_{(1-759)}$w/oST

The Gibson Assembly Master Mix includes three different enzymes in a single buffer: The exonuclease that creates single-stranded 3´ overhangs, allowing the annealing of fragments that share complementarity at one end. DNA polymerase fills in gaps within each annealed fragment. The DNA ligase seals nicks in the assembled DNA resulting in a double-stranded DNA molecule (Gibson et al., 2010; Gibson et al., 2009).

The gene insert was amplified by PCR with oligonucleotides containing an overlapping complementarity with the vector. The vector was linearized by restriction digestion. The PCR fragment of the insert and the linearized vector was mixed in 0.5:1.0 pmol ratio. To that 10µl of Gibson Assembly® Master mix was added, the total volume made up to 20µl. The reaction was incubated at 50°C for 1hour. Two microliters of this reaction was used for *E.coli* transformation.

4.2.1.4. pET 21b constructs for protein expression in E. coli

pET 21b-His$_6$-Toc159 $_{1-727}$ (Toc159A wt) was a gift from Dr. Matthew D. Smith, Wilfrid Laurier University, Canada (Richardson et al., 2009).

pET21b-His$_6$-Toc159 $_{1-759}$ (Toc159A w/o ST) was engineered using the Gibson Assembly® technique and a PCR amplified fragment of 1-759 amino acids derived either from the full-length synthetic wild type Toc159 cDNA or atToc159 cDNA with all of its serine and threonine codons replaced by alanine codons (produced by GeneCust, Luxembourg) (Fig. 23). The pET21b plasmid backbone was amplified for Gibson assembly® using GpF 5'-TAA GTC GAC AAG CTT GCG G-3' and GpR 5'-GTG GTG GTG GTG GTG GTG-3', the insert was amplified using GtF 5'- AGC CAC CAC CAC CAC CAC CAC GAT GCT AAA GCT GCC CCT GAA C-3' and GtR 5'-GCA AGC TTG TCG ACT TAA CGA TTC GCC CGG GGT GC-3' at an annealing temperature of 72°C.

4.2.1.5 pEG101-3HA-Toc1591-759-YFP (pEG101-Toc159A-YFP) for in planta expression

The A-domain of Toc159 was derived from the full-length synthetic wild type atToc159 cDNA, which was produced by GeneCust, Luxembourg (Fig. 23). The region of interest (coding for amino acids 1-759) was amplified by PCR using forward ACAAGTTTGTACAAAAAAGCAGGCTATGGCATACCCGTACGATGTTCC and reverse ACCACTTTGTACAAGAAAGCTGGGTAACGATTCGCCCGGGGTGCTG containing attB sites and gene specific region and cloned into pDONR221, an entry vector for gateway cloning. The entry vector was used to recombine the Toc159 $_{1-759}$ coding sequence into

the pEG101 destination vector resulting in a C-terminal YFP tag upon *in planta* expression.

4.2.1.6. DNA Constructs used for coupled in vitro transcription/translation

pET21d-prRBCS, was a gift from Dr. Birgit Agne, Martin-Luther-Universität Halle-Wittenberg, Germany. pET21d-prE1α-DHFR$_{His}$ was kindly donated by Prof. Danny Schnell, University of Massachusetts, USA (Inoue et al., 2010) pBluescript-prTic40 was a gift from Dr. Hosu-Min Li, Academia Sinica, Taipei, Taiwan (Teng et al., 2012).

4.2.1.7. Sequence verification of DNA constructs

All DNA constructs were sequenced for confirmation by Microsynth GmbH. The results are annexed (Appendix.1).

4.2.2. Bacterial transformation

4.2.2.1. Transformation of chemically competent E. coli

All the transformations of chemically competent *E. coli* were performed according to the transformation protocol provided by Invitrogen.

4.2.2.2. Transformation of electrocompetent Agrobacterium tumefaciens C58

An aliquot of competent C58 (100µl) was thawed on ice, 150ng of plasmid was added to the cells and mixed by tapping the tube. The mix was then transferred to a cold electroporation cuvette. Using the settings for bacteria in the Micropulser ™, Bio-Rad (Electroporator) i.e. resistance: 200 ohms, capacitance: 25 µF, Voltage: 2.5 kV the Agrobacterium cells were electroporated. The cells were then recovered by adding 1ml of LB to the cuvette and regenerated at 28°C for 1-2hrs in the dark. The transformed cells were plated on to a solid LB plate with appropriate antibiotic along with Rifampicin. After 2-3 days of incubation at 28°C in the dark the colonies were selected and further analyzed.

4.2.3. Bacterial expression and purification

4.2.3.1. Bacterial expression of pET21b constructs expression and protein purification on ÄKTAPrime®

pET21b constructs were transformed into *E.coli* BL21 (DE3) cells for protein expression. Cells were grown to a density of 0.8 OD_{600}. Protein expression was induced by the addition of 1 mM IPTG and expressed for 3 hrs prior to harvest by centrifugation at 1500g for 30 min @4°C. The cells were lysed using a French press in a buffer containing 50mM Tris-HCl, 300mM NaCl and 5mM Imidazol. The overexpressed protein was purified using Ni-NTA column on a FPLC, ÄKTAPrime®. The protein was eluted with 300mM imidazole containing buffer the fractions with recombinant protein was dialyzed against 20mM Piperazine and 50mM NaCl pH 5.5. The dialyzed Ni-NTA eluate was repurified on DEAE ion exchange column using FPLC-ÄKTAPrime®. Then concentrated using Amicon® 35kDa filters according to manufacturers recommendation.

4.2.4. Plant protein extraction by grinding or steel ball mill for TAP tag pull down

Arabidopsis seedlings with or without chemical crosslinking were frozen in liquid N_2, steel containers with steel balls were also chilled in liquid N_2 prior to sample application. The samples were milled in a vibrational shaker for three cycles of 1 min "on" and 1 min "off" in liquid N_2. The powder was directly collected in a 50mL falcon tube with grinding buffer (100mM Tris HCl pH 7.5, 200mM NaCl, 0.2% PIC, 1mM PMSF and 5mM NaF) with 0.75% TritonX-100. The samples were thawed in grinding buffer for 30 min. Larger plant debris were removed by centrifugation at 1500g for 10 min at 4°C. The supernatant was centrifuged at 100,000g for an hour at 4°C to obtain a clear extract (input).

4.2.5. Protein analysis by PAGE and Western blot

4.2.5.1. SDS PAGE and Western blotting

The SDS-PAGE minigels used for protein analysis were cast using the Bio-Rad system (Bio-Rad Laboratories, USA). The running gel percentage was between 10-12% according to experimental requirements. The composition of the resolving gel was 10-12% (w/v) acrylamide/bisacrylamide 37.5:1 (Bio-Rad), 0.4M Tris-HCl pH8.8, 0.1% (w/v) SDS, 0.05% (v/v) TEMED (Bio-Rad) and 0.08% (w/v) APS (Bio-Rad). The stacking gel had 4% (w/v) acrylamide/bisacrylamide 37.5:1 (Bio-Rad), 0.06M Tris-HCl pH6.8, 0.1% (w/v)

SDS, 0.1% (v/v) TEMED (Bio-Rad) and 0.08% (w/v) APS (Bio-Rad). The protein samples were dissolved in 1x Lämmli buffer (0.04M Tris-HCl pH6.8, 2% (w/v) SDS, 10% (v/v) Glycerol, 0.1% (w/v) Bromophenol blue with or without 0.1M DTT and heat denatured at 65°C for 15 min prior to loading on to the SDS-PAGE.

After resolving the protein samples by electrophoresis the gel was either stained with Coomassie blue according to standard protocols (Sambrook and Russell, 2001) and dried, or used for Western blot transfer to Protran® nitrocellulose membrane (GE Lifesciences) according to standard protocols (Sambrook and Russell, 2001).

The Western blot membrane was stained with Amido black according to standard protocol (Sambrook and Russell, 2001), destained (in 40% (v/v) methanol and 10% (v/v) acetic acid) and scanned. The membrane was blocked with 5% (w/v) skimmed milk powder dissolved either in 1xTBS or 1xPBS buffer according to the antibody supplier's recommendation. Subsequently the membrane was incubated with appropriate concentrations of primary antibody in 5% skimmed milk TBS or PBS for 1 hour (except for anti-Toc159A and anti-CBP, which were incubated over night) After washing with 1xTBS or PBS for 30 min, the membrane was incubated with and anti-rabbit IgG secondary antibody coupled to horseradish peroxidase in 5% skimmed milk TBS or PBS for 30 min. The blots washed for 3 x10 minutes with xx, the membrane was treated with a chemoluminescence reagent mixture (0.1M Tris-HCl pH 8.5, 0.2 mM p-coumaric acid, 1.25mM luminol (3-aminophthalhydrazide) along with 0.009% H_2O_2, the chemiluminescent signals were detected by exposing to X-ray films (Eastman Kodak Company, USA).

4.2.5.2. BN PAGE and Western blot analysis

The 4-16% gradient gels (14 x 16 x 1 cm) used for BN-PAGE (Blue Native-Polyacrylamide gel electrophoresis) were cast using a Hoefer SE600 Ruby system from Amersham Biosciences, USA. The gel composition and protocols for Western blotting were adapted from Kikuchi et al., 2009. Total plant ground with liq.N_2 equivalent to 60µg chlorophyll were solubilized in BN-PAGE sample buffer (50 mM BisTris-HCl, pH 7.0, 500mM 6-amino-n-caproic acid, 10% (w/v) glycerol and 1% digitonin, containing 0.1% protease inhibitor cocktail). The mixture was left on ice for 20 min, non-solubilized material was removed by centrifugation at 25,000xg for 30 min at 4°C, the supernatant was mixed with 5% G-250 (5% Coomassie G-250 in 50mM BisTris-HCl pH7, 500mM 6-amino-n-caproic acid) and loaded on to the gel.

The gel was run in 1x running buffer (50mM BisTris-HCl, 50mM Tricine pH 6.8) for overnight at constant 30V. The molecular mass marker was a mixture of Ferritin and Bovine serum albumin dissolved in 5% G-250.

For Western blotting the gel was denatured in 65mM Tric-HCl pH 6.8 buffer with 3.3% SDS and 4% β-mercaptoethanol by heating at 80°C for 30 min. The PVDF membrane was pre-soaked in 100% methanol and sandwiched with the denatured gel for overnight transfer in blotting buffer (100mM Tris, 192mM Glycine, 0.1% SDS and 20% methanol) at constant 40V. The membrane was fixed with fixing solution (40% methanol, 10% acetic acid) for x minutes, then destained in 100% methanol for x minutes and washed in 1xPBS. Thereafter the protocol described for SDS-PAGE Western blotting was followed.

4.2.5.3. Protein complex Immunodepletion with anti-CBP antibody

Ground plant tissue, with or without chemical crosslinking, equivalent to 80µg chlorophyll was solubilized in BN-PAGE sample buffer (50 mM BisTris-HCl, pH 7.0, 500mM 6-amino-n-caproic acid, 10% (w/v) glycerol and 1% digitonin, with 0.1% protease inhibitor cocktail). Incubated for 20 min on ice with frequent mixing by tapping, insoluble matter was removed by centrifugation at 25,000xg for 30 min at 4°C. Meanwhile 10µg of anti-Calmodulin Binding Protein (anti-CBP) antibodies were incubated with 10µL of ProtA-sepharose magnetic beads (Millipore, USA). The supernatant from the centrifugation was added to the anti-CBP-ProtA-sepharose magnetic beads mixture, allowed to bind on a rotary wheel for an hour. The unbound protein was recovered and passed through 0.2µm filter to remove any sepharose beads. The filtrates were subjected to BN-PAGE.

4.2.6. Arabidopsis cultivation and transformation

4.2.6.1. Growing Arabidopsis thaliana on Murashige and Skoog medium

Seeds were surface sterilized with 70% (w/v) Ethanol and 0.05% Triton for 5 min followed by 100% Ethanol for 5 min. The seeds were spread on plates consisting of ½ MS medium with 1% phytoagar and 0.8% Sucrose. Germination was synchronized by exposing the plates to 4°C for 2 days in dark. For all the experiments except for seed propagation the plants were grown under short day conditions (8h light/16h dark) at a light intensity of 120µE m^{-2} s^{-2}. For selection of transformed plants the ½ MS plates were supplemented with 30µg/ml of phosphinothricin (BASTA; Duchefa).

4.2.6.2. Arabidopsis protoplast transient transfection

Transient transformation of protoplasts with pEG101-3HA-Toc159A-YFP was done using polyethylene glycol method as described (Jin et al., 2001) but the cellulase and macerozyme (Serva) concentrations were changed to 1% and 0.25% (w/v) respectively.

4.2.6.3. Stable Arabidopsis transformation using Agrobacterium tumefaciens

Arabidopsis wild type plants were transformed with pEG101-Toc159A-YFP using floral dip method as described (Clough & Bent, 1998). The transformed plants were selected on plates containing phosphinothricin (see 4.2.6.1). The selected lines were then characterized by PCR and western blotting against HA tag to verify the protein expression.

4.2.6.4. Confocal laser scanning microscopy imaging

Fluorescence signals in transiently transformed protoplasts as well as stably transformed plants were observed by confocal laser scanning microscopy using a LEICA TCS SP2 microscope (LEICA Microsystems) to detect YFP signals. Excitiation wavelength was 514nm. Fluorescence emission was recorded at 524nm.

4.2.7. Arabidopsis protoplast/chloroplast extraction

Arabidopsis protoplasts and chloroplasts from 3 weeks old seedlings were isolated according to the enzymatic digestion protocol described by Smith et al..(Smith et al., 2003). Only the ariel tissues of Arabidopsis seedlings grown on ½ MS plates were cut and washed in distilled water. Then washed seedlings were chopped with a scalpel for 30 seconds while immersed in digestion buffer without enzymes (400 mM sorbitol, 20mM MES-KOH pH 5.2, 0.5 mM $CaCl_2$), after leaving the plants to stand on ice for 5 mins the buffer was poured off and washing was repeated for one more time before adding the digestion buffer with cellulase and macerozyme (from *ONOZUKA*; 3% and 0.75% respectively). Seedlings were left in shade for minimum of 4 hours before gently agitating them to release the protoplasts. Isolated protoplasts were washed with protoplast resuspension buffer (digestion buffer pH 6.0), clean protoplast pellet was resuspended in breakage buffer (300 mM sorbitol, 20mM Tricine-KOH pH 8.4, 5 mM EDTA, 5 mM EGTA, 10 mM $NaHCO_3$, 0.1% BSA) and then passed through a modified syringe set up with two layers of mira-cloth (23µm and 18µm). Intact chloroplasts were

separated on a Percoll® gradient, washed and resuspended in HEPES-Sorbitol buffer (50mM HEPES-KOH, 330mM Sorbitol pH 8.0). Further, experimentation with isolated chloroplasts such as *in vitro* protein import and thermolysin treatment was also performed as mentioned in the protocol.

For IgG pull down of soluble and membrane fraction complexes, the protoplasts either treated with or without DSP chemical crosslinking were broken by passing them through a nylon syringe set up described in the protocol and the broken mixture loaded onto Percoll® gradient. After centrifugation the top layer with broken chloroplasts along with all soluble contents of the cytosol were taken as soluble fraction, and the intact chloroplasts were washed and solubilized with 1% DDM to get pellet or membrane fraction.

4.2.8. *In vitro* methods

4.2.8.1. In vitro preprotein translation

Standard *in vitro* preprotein translation

A standard *in vitro* transcription/translation reaction was performed with 40µl of reticulocyte lysate master mix (TNT®Quick coupled Transcription/Translation system, Promega, USA) 2µl of ^{35}S-Methionine (10mCi/ml) with plasmid DNA template 0.5µg and water to a total volume of 50µl. The mixture was incubated at 30°C for 90 min and stored in -20°C until use.

In vitro preprotein translation in the presence of phosphorylated Toc159A

Eight microliters of phosphorylated recombinant Toc159A protein was added to a standard *in vitro* transcription/translation reaction and incubated for 90 min at 30°C, stored in -20°C until use.

4.2.8.2. Phosphorylation of recombinant proteins in vitro

The recombinant Toc159A, either Toc159A wt or Toc159A w/o ST, was subjected phosphorylation *in vitro* as described (Agne et al., 2010) using 100µg of recombinant protein (10-20µg/µl) per reaction and non-radioactive ATP.

4.2.9. *In vivo* methods

4.2.9.1. Chemical crosslinking using DSP in Arabidopsis protoplasts

For *in vivo* crosslinking of proteins interacting with both in the cytosol and chloroplast localized Toc159A Dithiobis-[succinimidylpropionate] DSP, a membrane permeable reversible crosslinker, was used on Arabidopsis protoplasts. 50mM DSP crosslinker was prepared in 1xPBS buffer pH 7.4. Isolated Arabidopsis protoplasts (see 4.2.7) were treated with DSP crosslinker at a final concentration of 5mM DSP in. The mixture was incubated in room temperature for 45 min. After removing DSP from the mixture by centrifuging at 100g for 5 min, any excess DSP was quenched by the addition of 200mM Tris pH 7.4 for 10 min at room temperature; then the protoplasts were washed in HEPES-Sorbitol buffer pH 8.0 (50mM HEPES-KOH, 330mM Sorbitol). Subsequently the protoplasts were ruptured, the cytosolic or soluble fraction was separated from membrane or pellet fraction by centrifugation of broken protoplasts at 100,000xg. The pellet was washed once in HEPES-Sorbitol buffer then solubilized with 1% DDM (n-Dodecyl β-D-Maltopyranoside) to prepare the membrane fraction, both the fractions were chloroform methanol precipitated, resuspended in appropriate sample buffer for SDS-PAGE-immuno blot analysis.

4.2.9.2. Vacuum infiltration of DSP into Arabidopsis seedlings

To facilitate permeation by the reversible chemical crosslinker DSP, vacuum infiltration of intact seedlings method was carried out. Three weeks old Arabidopsis seedlings in a Falcon tube were immersed in 5mM DSP solution in 1:6 ratio(w/v). Vacuum was applied using a SpeedVac (Eppendorf) for 15min; the release of the vacuum aided the penetration of DSP into the intact seedlings. Seedlings were incubated in the DSP crosslinker solution for another 30 min. The solution was removed by decanting and the seedlings were washed with 1xPBS buffer prior to steel mill grinding in grinding buffer (100mM Tris HCl pH 7.5, 200mM NaCl, 0.2% PIC, 1mMPMSF and 5mM NaF). The insoluble plant material was removed by centrifugation at 1500g for 5 min. The supernatant was subjected to ultracentrifugation at 100,000g in 4°C for 1 hour, and the supernatant of this step was considered the soluble fraction. The pellet from the 100'000g centrifugation step was resuspended in grinding buffer containing 1% DDM and incubated for 30 min prior to centrifugation again at 100,000g. This time the supernatant was considered as the solubilized membrane fraction. Proteins from the fractions were extracted using the chloroform:methanol protocol and analysed by SDS-PAGE followed by Western blotting.

4.2.9.3. Formaldehyde crosslinking and immunoprecipitation of Toc159A-TAP complex.

Vacuum infiltration of formaldehyde for in vivo crosslinking

Three weeks old Arabidopsis seedlings were immersed in a 4% formaldehyde solution prepared from paraformaldehyde in 1xPBS buffer (or as mentioned) at a 1:5 ratio (w/v). Vacuum was applied using a SpeedVac (Eppendorf) for 15 min, the removal of the vacuum aided the penetration of the formaldehyde into the intact seedlings. Seedlings were left in the formaldehyde solution for another 10 min. The solution was decanted and unreacted formaldehyde was quenched using 1.25M Glycine (the free amine groups reacted with reactive hydrogen of formaldehyde). The seedlings were washed twice with 1xPBS buffer. The seedlings were dried by patting with paper towels, frozen in liquid N_2 and stored in -80°C until used grinding in a steel ball mill (see 2.2.4).

IgG pull down of Toc159A-TAP

Total protein extract (Input) clarified at 100,000g for 1hr was allowed to bind with Hs IgG-sepharose beads equilibrated with grinding buffer (50 mM HEPES KOH, pH 7.5, 100 mM NaCl, 0.5% polyvinylpyrrolidone, 5 mM NaF, 1 mM DTT, 0.1% [v/v] protease inhibitor cocktail, and 1 mM phenylmethylsulfonyl fluoride), over night at 4°C. The Hs IgG-sepharose was filled into a polypropylene column. The clarified total protein extract was allowed to pass through the polypropylene column. The unbound fraction was collected as flowthrough fraction for analysis. The Hs IgG-sepharose was washed with grinding buffer with or without detergent using 3-4 times the volume of the input extracted. Proteins bound to the column were eluted with 100µl TEV protease (200U/mL in TEV elution buffer) according to the manufacturer's guideline, 2 x 200µl of 200mM Glycine pH 3.0 or 2% SDS. Protein was extracted using the chloroform: methanol protocol prior to SDS-PAGE separation and Western blot analysis or processed for Mass-spectrometry analysis.

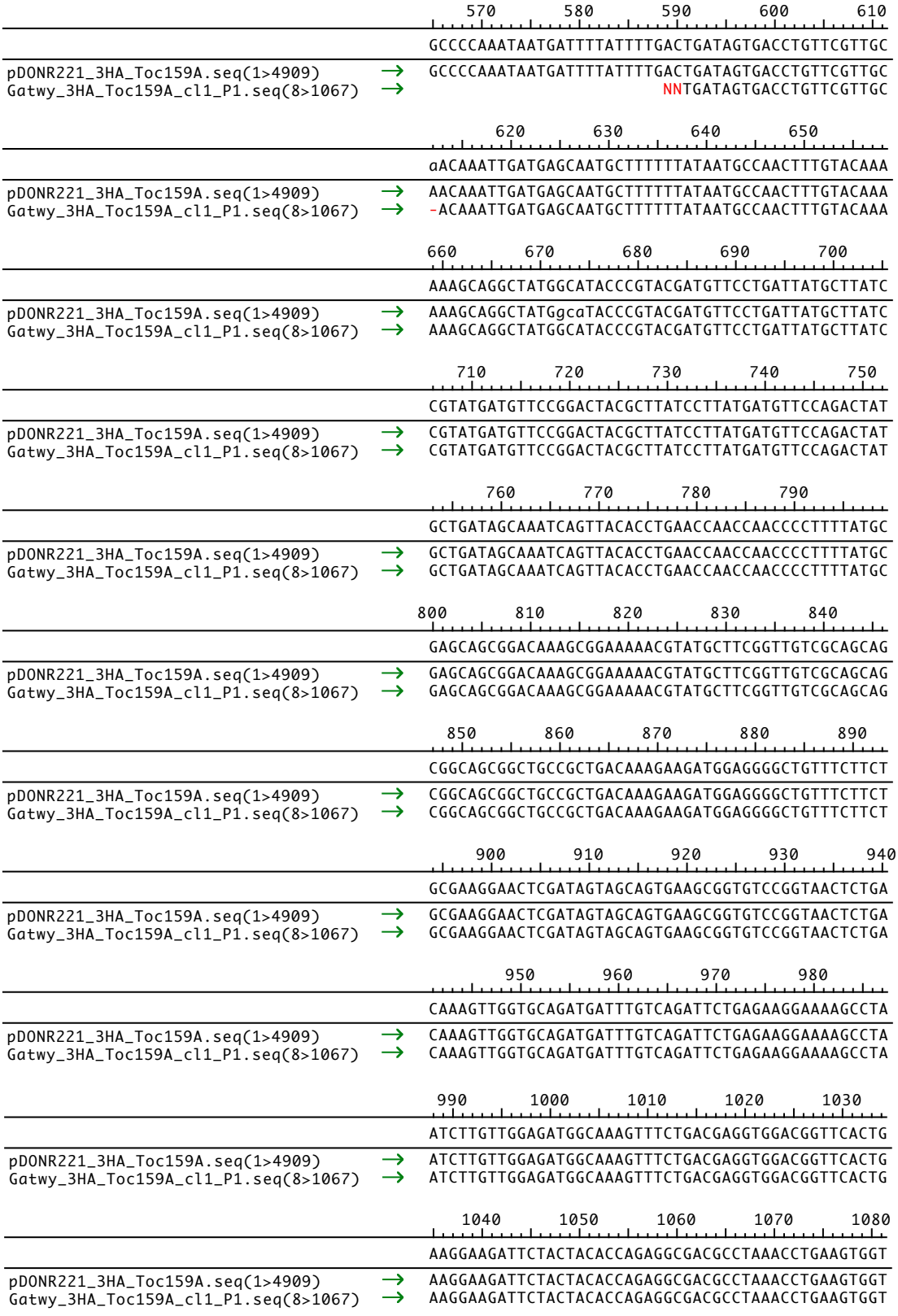
4.2.9.4. Mass spectrometry analysis

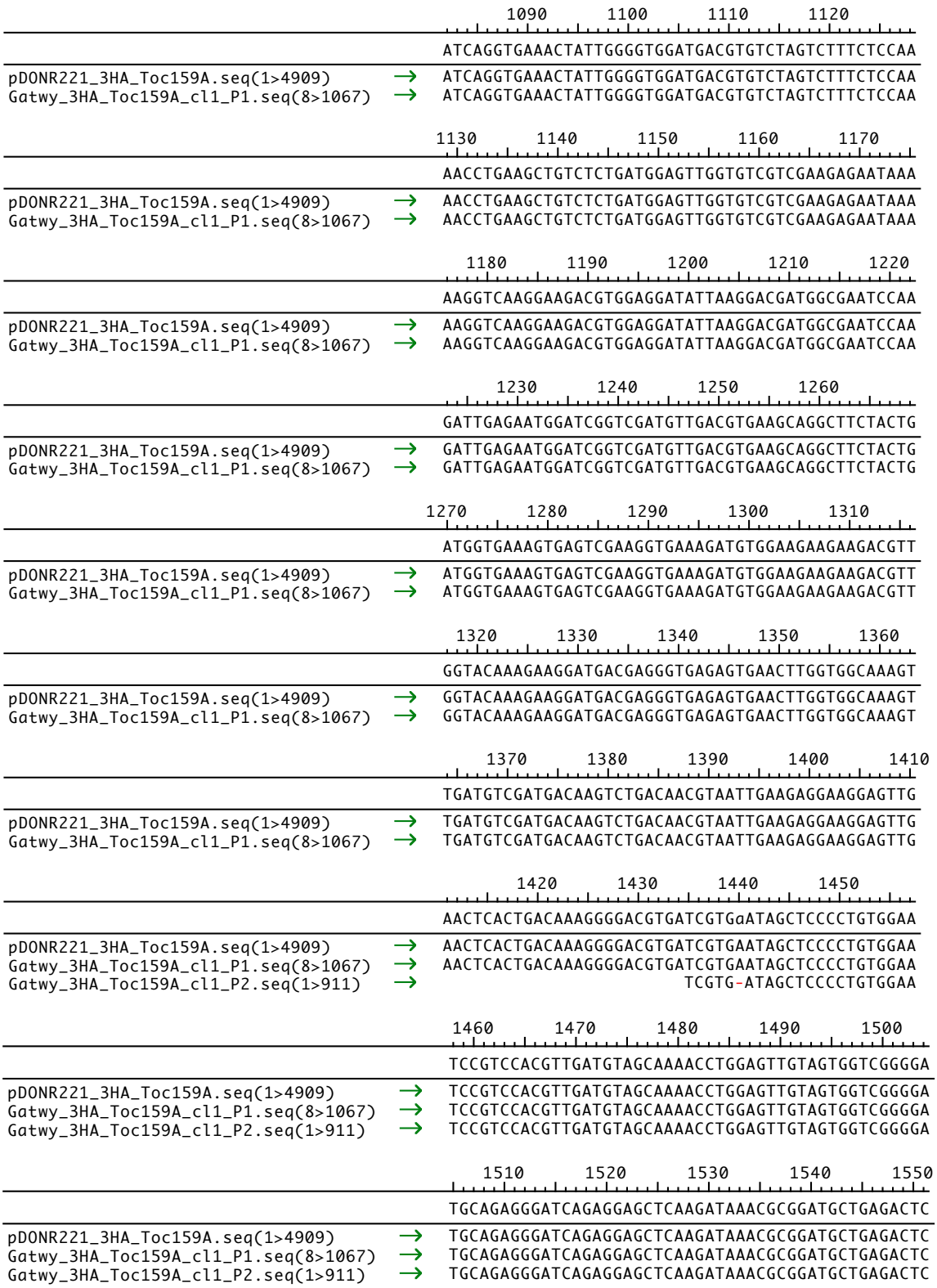
Only part of the procedure was performed in University of Neuchatel, the dried peptides were sent to Dr. Birgit Agne, Martin-Luther-Universität Halle-Wittenberg, Germany. The protocol was followed with modifications from (Agne et al., 2010).

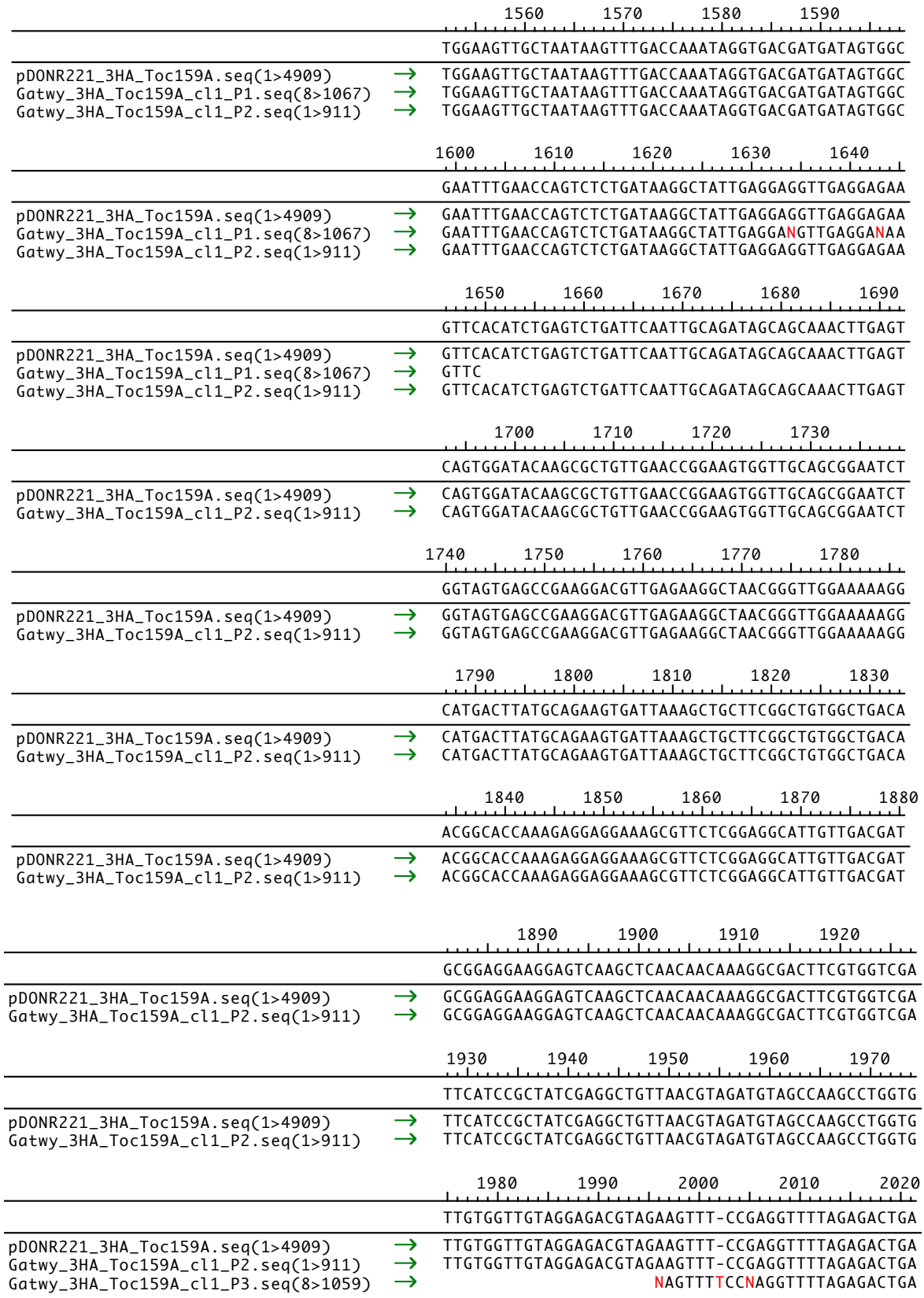
Proteins separated on 4% to 12% SDS gels were stained using SyproRuby®. Each selected gel piece was sliced into small pieces. In-gel tryptic digestion was performed according to a modified protocol from Shevchenko et al. (Shevchenko et al., 1996; 2006). Prior to mass spectrometric analysis, peptide samples were desalted using SepPak cartridges (Waters). Dried peptides were resuspended in 3% (v/v) acetonitrile (ACN), 0.2% (v/v) formic acid and analyzed on a LTQ FT-ICR mass spectrometer (Thermo Fisher Scientific) coupled with an Eksigent nanoliquid chromatography system (Eksigent Technologies). Peptide mixtures were loaded onto laboratory-made capillary columns (75 mm i.d.; BGB Analytik), 8 cm length, packed with Magic C18 AQ beads (3 mm, 100 A° ; Michrom BioResources). Peptides were eluted from the column by an increased ACN concentration in the mobile phase from 5% (v/v) ACN, 0.2% (v/v) formic acid to 40% (v/v) ACN, 0.2% (v/v) formic acid over 74 min, followed by a 10 min wash step at 5% (v/v) ACN, 0.2% (v/v) formic acid. Peptide ions were detected in a survey scan from 300 to 1,600 atomic mass units at 100,000 full width half maximum nominal resolution followed by three data-dependent tandem mass spectrometry scans (isolation width 2 atomic mass units, relative collision energy 35%, dynamic exclusion enabled, repeat count 1, followed by peak exclusion for 2 min). Tandem mass spectrometry spectra were searched using Mascot 2.1.04 (Matrix Science) against The Arabidopsis Information Resource 8 protein database supplemented with contaminants. The search parameters were as follows: requirement for tryptic ends, one missed cleavage allowed, mass tolerance = 5 ppm.

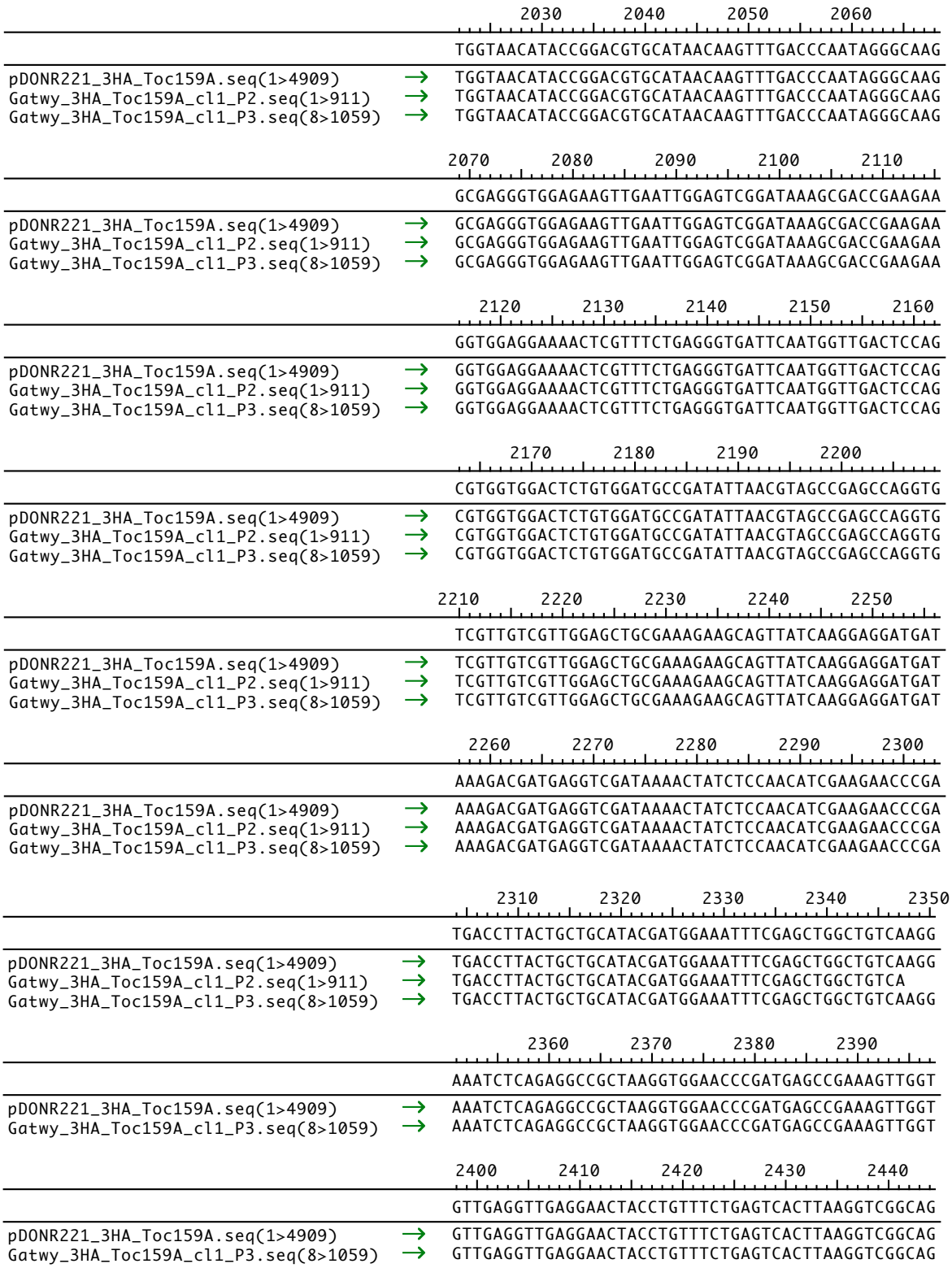
6. Appendix

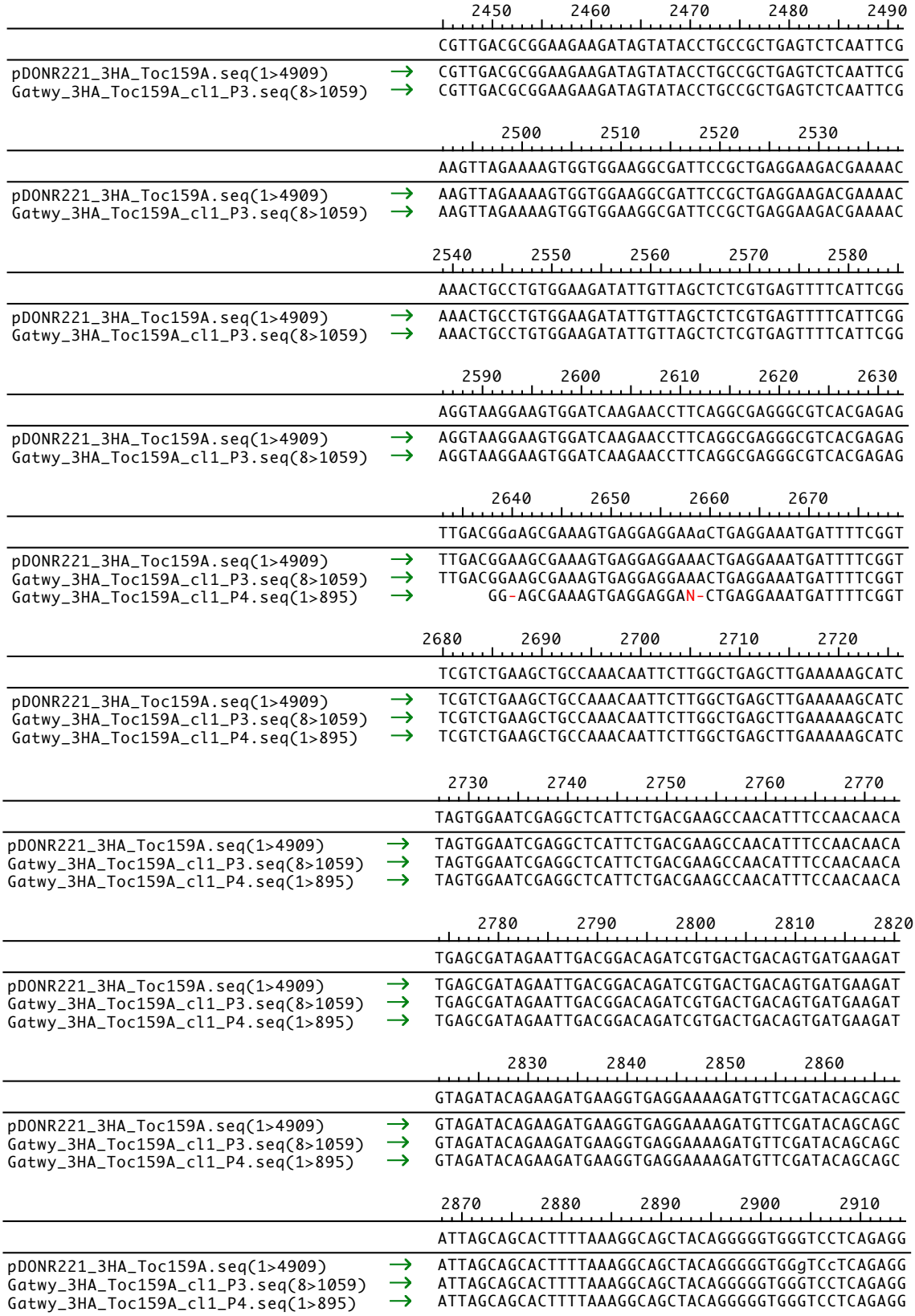
pEG101-3HA-TOC159A-YFP sequence result

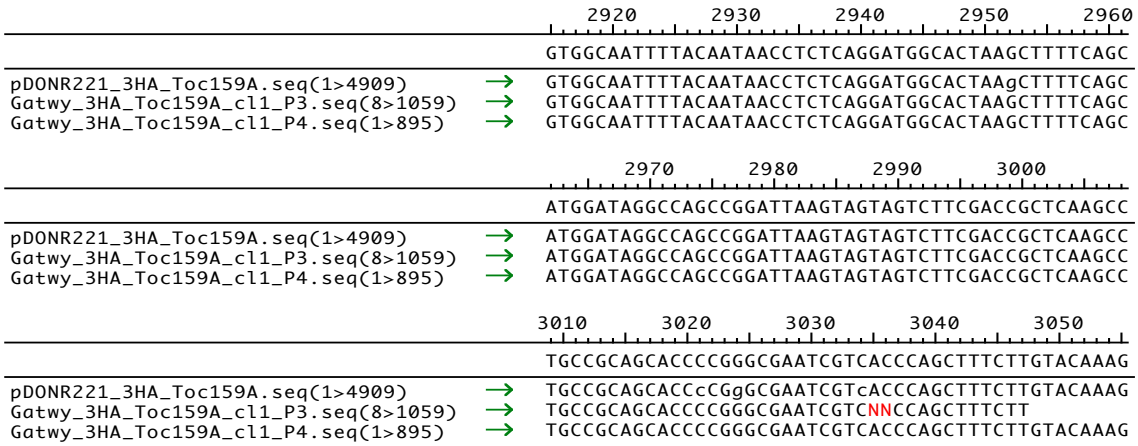



1090 1100 1110 1120
ATCAGGTGAAACTATTGGGGTGGATGACGTGTCTAGTCTTTCTCCAA
pDONR221_3HA_Toc159A.seq(1>4909) → ATCAGGTGAAACTATTGGGGTGGATGACGTGTCTAGTCTTTCTCCAA
Gatwy_3HA_Toc159A_cl1_P1.seq(8>1067) → ATCAGGTGAAACTATTGGGGTGGATGACGTGTCTAGTCTTTCTCCAA
1130 1140 1150 1160 1170
AACCTGAAGCTGTCTCTGATGGAGTTGGTGTCGTCGAAGAGAATAAA
pDONR221_3HA_Toc159A.seq(1>4909) → AACCTGAAGCTGTCTCTGATGGAGTTGGTGTCGTCGAAGAGAATAAA
Gatwy_3HA_Toc159A_cl1_P1.seq(8>1067) → AACCTGAAGCTGTCTCTGATGGAGTTGGTGTCGTCGAAGAGAATAAA
1180 1190 1200 1210 1220
AAGGTCAAGGAAGACGTGGAGGATATTAAGGACGATGGCGAATCCAA
pDONR221_3HA_Toc159A.seq(1>4909) → AAGGTCAAGGAAGACGTGGAGGATATTAAGGACGATGGCGAATCCAA
Gatwy_3HA_Toc159A_cl1_P1.seq(8>1067) → AAGGTCAAGGAAGACGTGGAGGATATTAAGGACGATGGCGAATCCAA
1230 1240 1250 1260
GATTGAGAATGGATCGGTCGATGTTGACGTGAAGCAGGCTTCTACTG
pDONR221_3HA_Toc159A.seq(1>4909) → GATTGAGAATGGATCGGTCGATGTTGACGTGAAGCAGGCTTCTACTG
Gatwy_3HA_Toc159A_cl1_P1.seq(8>1067) → GATTGAGAATGGATCGGTCGATGTTGACGTGAAGCAGGCTTCTACTG
1270 1280 1290 1300 1310
ATGGTGAAAGTGAGTCGAAGGTGAAAGATGTGGAAGAAGAAGACGTT
pDONR221_3HA_Toc159A.seq(1>4909) → ATGGTGAAAGTGAGTCGAAGGTGAAAGATGTGGAAGAAGAAGACGTT
Gatwy_3HA_Toc159A_cl1_P1.seq(8>1067) → ATGGTGAAAGTGAGTCGAAGGTGAAAGATGTGGAAGAAGAAGACGTT
1320 1330 1340 1350 1360
GGTACAAAGAAGGATGACGAGGGTGAGAGTGAACTTGGTGGCAAAGT
pDONR221_3HA_Toc159A.seq(1>4909) → GGTACAAAGAAGGATGACGAGGGTGAGAGTGAACTTGGTGGCAAAGT
Gatwy_3HA_Toc159A_cl1_P1.seq(8>1067) → GGTACAAAGAAGGATGACGAGGGTGAGAGTGAACTTGGTGGCAAAGT
1370 1380 1390 1400 1410
TGATGTCGATGACAAGTCTGACAACGTAATTGAAGAGGAAGGAGTTG
pDONR221_3HA_Toc159A.seq(1>4909) → TGATGTCGATGACAAGTCTGACAACGTAATTGAAGAGGAAGGAGTTG
Gatwy_3HA_Toc159A_cl1_P1.seq(8>1067) → TGATGTCGATGACAAGTCTGACAACGTAATTGAAGAGGAAGGAGTTG
1420 1430 1440 1450
AACTCACTGACAAAGGGGACGTGATCGTGaATAGCTCCCCTGTGGAA
pDONR221_3HA_Toc159A.seq(1>4909) → AACTCACTGACAAAGGGGACGTGATCGTGAATAGCTCCCCTGTGGAA
Gatwy_3HA_Toc159A_cl1_P1.seq(8>1067) → AACTCACTGACAAAGGGGACGTGATCGTGAATAGCTCCCCTGTGGAA
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → TCGTG-ATAGCTCCCCTGTGGAA
1460 1470 1480 1490 1500
TCCGTCCACGTTGATGTAGCAAAACCTGGAGTTGTAGTGGTCGGGGA
pDONR221_3HA_Toc159A.seq(1>4909) → TCCGTCCACGTTGATGTAGCAAAACCTGGAGTTGTAGTGGTCGGGGA
Gatwy_3HA_Toc159A_cl1_P1.seq(8>1067) → TCCGTCCACGTTGATGTAGCAAAACCTGGAGTTGTAGTGGTCGGGGA
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → TCCGTCCACGTTGATGTAGCAAAACCTGGAGTTGTAGTGGTCGGGGA
1510 1520 1530 1540 1550
TGCAGAGGGATCAGAGGAGCTCAAGATAAACGCGGATGCTGAGACTC
pDONR221_3HA_Toc159A.seq(1>4909) → TGCAGAGGGATCAGAGGAGCTCAAGATAAACGCGGATGCTGAGACTC
Gatwy_3HA_Toc159A_cl1_P1.seq(8>1067) → TGCAGAGGGATCAGAGGAGCTCAAGATAAACGCGGATGCTGAGACTC
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → TGCAGAGGGATCAGAGGAGCTCAAGATAAACGCGGATGCTGAGACTC


1560 1570 1580 1590
TGGAAGTTGCTAATAAGTTTGACCAAATAGGTGACGATGATAGTGGC
pDONR221_3HA_Toc159A.seq(1>4909) → TGGAAGTTGCTAATAAGTTTGACCAAATAGGTGACGATGATAGTGGC
Gatwy_3HA_Toc159A_cl1_P1.seq(8>1067) → TGGAAGTTGCTAATAAGTTTGACCAAATAGGTGACGATGATAGTGGC
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → TGGAAGTTGCTAATAAGTTTGACCAAATAGGTGACGATGATAGTGGC
1600 1610 1620 1630 1640
GAATTTGAACCAGTCTCTGATAAGGCTATTGAGGAGGTTGAGGAGAA
pDONR221_3HA_Toc159A.seq(1>4909) → GAATTTGAACCAGTCTCTGATAAGGCTATTGAGGAGGTTGAGGAGAA
Gatwy_3HA_Toc159A_cl1_P1.seq(8>1067) → GAATTTGAACCAGTCTCTGATAAGGCTATTGAGGANGTTGAGGANAA
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → GAATTTGAACCAGTCTCTGATAAGGCTATTGAGGAGGTTGAGGAGAA
1650 1660 1670 1680 1690
GTTCACATCTGAGTCTGATTCAATTGCAGATAGCAGCAAACTTGAGT
pDONR221_3HA_Toc159A.seq(1>4909) → GTTCACATCTGAGTCTGATTCAATTGCAGATAGCAGCAAACTTGAGT
Gatwy_3HA_Toc159A_cl1_P1.seq(8>1067) → GTTC
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → GTTCACATCTGAGTCTGATTCAATTGCAGATAGCAGCAAACTTGAGT
1700 1710 1720 1730
CAGTGGATACAAGCGCTGTTGAACCGGAAGTGGTTGCAGCGGAATCT
pDONR221_3HA_Toc159A.seq(1>4909) → CAGTGGATACAAGCGCTGTTGAACCGGAAGTGGTTGCAGCGGAATCT
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → CAGTGGATACAAGCGCTGTTGAACCGGAAGTGGTTGCAGCGGAATCT
1740 1750 1760 1770 1780
GGTAGTGAGCCGAAGGACGTTGAGAAGGCTAACGGGTTGGAAAAAGG
pDONR221_3HA_Toc159A.seq(1>4909) → GGTAGTGAGCCGAAGGACGTTGAGAAGGCTAACGGGTTGGAAAAAGG
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → GGTAGTGAGCCGAAGGACGTTGAGAAGGCTAACGGGTTGGAAAAAGG
1790 1800 1810 1820 1830
CATGACTTATGCAGAAGTGATTAAAGCTGCTTCGGCTGTGGCTGACA
pDONR221_3HA_Toc159A.seq(1>4909) → CATGACTTATGCAGAAGTGATTAAAGCTGCTTCGGCTGTGGCTGACA
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → CATGACTTATGCAGAAGTGATTAAAGCTGCTTCGGCTGTGGCTGACA
1840 1850 1860 1870 1880
ACGGCACCAAAGAGGAGGAAAGCGTTCTCGGAGGCATTGTTGACGAT
pDONR221_3HA_Toc159A.seq(1>4909) → ACGGCACCAAAGAGGAGGAAAGCGTTCTCGGAGGCATTGTTGACGAT
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → ACGGCACCAAAGAGGAGGAAAGCGTTCTCGGAGGCATTGTTGACGAT
1890 1900 1910 1920
GCGGAGGAAGGAGTCAAGCTCAACAACAAAGGCGACTTCGTGGTCGA
pDONR221_3HA_Toc159A.seq(1>4909) → GCGGAGGAAGGAGTCAAGCTCAACAACAAAGGCGACTTCGTGGTCGA
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → GCGGAGGAAGGAGTCAAGCTCAACAACAAAGGCGACTTCGTGGTCGA
1930 1940 1950 1960 1970
TTCATCCGCTATCGAGGCTGTTAACGTAGATGTAGCCAAGCCTGGTG
pDONR221_3HA_Toc159A.seq(1>4909) → TTCATCCGCTATCGAGGCTGTTAACGTAGATGTAGCCAAGCCTGGTG
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → TTCATCCGCTATCGAGGCTGTTAACGTAGATGTAGCCAAGCCTGGTG
1980 1990 2000 2010 2020
TTGTGGTTGTAGGAGACGTAGAAGTTT-CCGAGGTTTTAGAGACTGA
pDONR221_3HA_Toc159A.seq(1>4909) → TTGTGGTTGTAGGAGACGTAGAAGTTT-CCGAGGTTTTAGAGACTGA
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → TTGTGGTTGTAGGAGACGTAGAAGTTT-CCGAGGTTTTAGAGACTGA
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → NAGTTTTCCNAGGTTTTAGAGACTGA

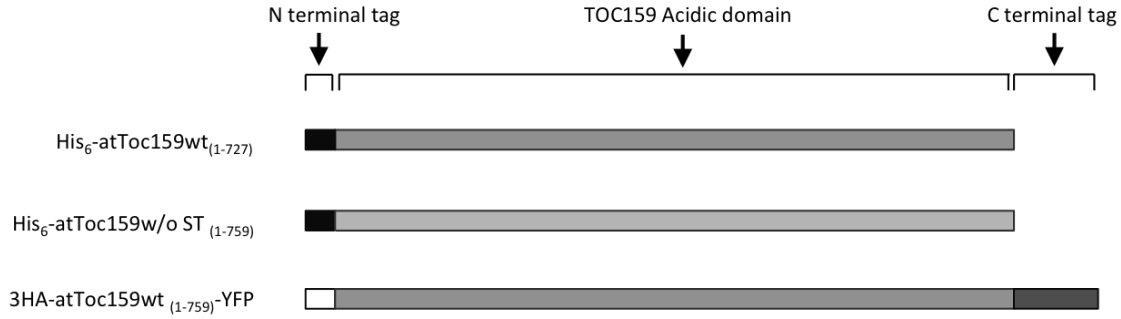

2030 2040 2050 2060
TGGTAACATACCGGACGTGCATAACAAGTTTGACCCAATAGGGCAAG
pDONR221_3HA_Toc159A.seq(1>4909) → TGGTAACATACCGGACGTGCATAACAAGTTTGACCCAATAGGGCAAG
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → TGGTAACATACCGGACGTGCATAACAAGTTTGACCCAATAGGGCAAG
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → TGGTAACATACCGGACGTGCATAACAAGTTTGACCCAATAGGGCAAG
2070 2080 2090 2100 2110
GCGAGGGTGGAGAAGTTGAATTGGAGTCGGATAAAGCGACCGAAGAA
pDONR221_3HA_Toc159A.seq(1>4909) → GCGAGGGTGGAGAAGTTGAATTGGAGTCGGATAAAGCGACCGAAGAA
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → GCGAGGGTGGAGAAGTTGAATTGGAGTCGGATAAAGCGACCGAAGAA
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → GCGAGGGTGGAGAAGTTGAATTGGAGTCGGATAAAGCGACCGAAGAA
2120 2130 2140 2150 2160
GGTGGAGGAAAACTCGTTTCTGAGGGTGATTCAATGGTTGACTCCAG
pDONR221_3HA_Toc159A.seq(1>4909) → GGTGGAGGAAAACTCGTTTCTGAGGGTGATTCAATGGTTGACTCCAG
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → GGTGGAGGAAAACTCGTTTCTGAGGGTGATTCAATGGTTGACTCCAG
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → GGTGGAGGAAAACTCGTTTCTGAGGGTGATTCAATGGTTGACTCCAG
2170 2180 2190 2200
CGTGGTGGACTCTGTGGATGCCGATATTAACGTAGCCGAGCCAGGTG
pDONR221_3HA_Toc159A.seq(1>4909) → CGTGGTGGACTCTGTGGATGCCGATATTAACGTAGCCGAGCCAGGTG
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → CGTGGTGGACTCTGTGGATGCCGATATTAACGTAGCCGAGCCAGGTG
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → CGTGGTGGACTCTGTGGATGCCGATATTAACGTAGCCGAGCCAGGTG
2210 2220 2230 2240 2250
TCGTTGTCGTTGGAGCTGCGAAAGAAGCAGTTATCAAGGAGGATGAT
pDONR221_3HA_Toc159A.seq(1>4909) → TCGTTGTCGTTGGAGCTGCGAAAGAAGCAGTTATCAAGGAGGATGAT
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → TCGTTGTCGTTGGAGCTGCGAAAGAAGCAGTTATCAAGGAGGATGAT
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → TCGTTGTCGTTGGAGCTGCGAAAGAAGCAGTTATCAAGGAGGATGAT
2260 2270 2280 2290 2300
AAAGACGATGAGGTCGATAAAACTATCTCCAACATCGAAGAACCCGA
pDONR221_3HA_Toc159A.seq(1>4909) → AAAGACGATGAGGTCGATAAAACTATCTCCAACATCGAAGAACCCGA
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → AAAGACGATGAGGTCGATAAAACTATCTCCAACATCGAAGAACCCGA
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → AAAGACGATGAGGTCGATAAAACTATCTCCAACATCGAAGAACCCGA
2310 2320 2330 2340 2350
TGACCTTACTGCTGCATACGATGGAAATTTCGAGCTGGCTGTCAAGG
pDONR221_3HA_Toc159A.seq(1>4909) → TGACCTTACTGCTGCATACGATGGAAATTTCGAGCTGGCTGTCAAGG
Gatwy_3HA_Toc159A_cl1_P2.seq(1>911) → TGACCTTACTGCTGCATACGATGGAAATTTCGAGCTGGCTGTCA
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → TGACCTTACTGCTGCATACGATGGAAATTTCGAGCTGGCTGTCAAGG
2360 2370 2380 2390
AAATCTCAGAGGCCGCTAAGGTGGAACCCGATGAGCCGAAAGTTGGT
pDONR221_3HA_Toc159A.seq(1>4909) → AAATCTCAGAGGCCGCTAAGGTGGAACCCGATGAGCCGAAAGTTGGT
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → AAATCTCAGAGGCCGCTAAGGTGGAACCCGATGAGCCGAAAGTTGGT
2400 2410 2420 2430 2440
GTTGAGGTTGAGGAACTACCTGTTTCTGAGTCACTTAAGGTCGGCAG
pDONR221_3HA_Toc159A.seq(1>4909) → GTTGAGGTTGAGGAACTACCTGTTTCTGAGTCACTTAAGGTCGGCAG
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → GTTGAGGTTGAGGAACTACCTGTTTCTGAGTCACTTAAGGTCGGCAG


2450 2460 2470 2480 2490
CGTTGACGCGGAAGAAGATAGTATACCTGCCGCTGAGTCTCAATTCG
pDONR221_3HA_Toc159A.seq(1>4909) → CGTTGACGCGGAAGAAGATAGTATACCTGCCGCTGAGTCTCAATTCG
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → CGTTGACGCGGAAGAAGATAGTATACCTGCCGCTGAGTCTCAATTCG
2500 2510 2520 2530
AAGTTAGAAAAGTGGTGGAAGGCGATTCCGCTGAGGAAGACGAAAAC
pDONR221_3HA_Toc159A.seq(1>4909) → AAGTTAGAAAAGTGGTGGAAGGCGATTCCGCTGAGGAAGACGAAAAC
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → AAGTTAGAAAAGTGGTGGAAGGCGATTCCGCTGAGGAAGACGAAAAC
2540 2550 2560 2570 2580
AAACTGCCTGTGGAAGATATTGTTAGCTCTCGTGAGTTTTCATTCGG
pDONR221_3HA_Toc159A.seq(1>4909) → AAACTGCCTGTGGAAGATATTGTTAGCTCTCGTGAGTTTTCATTCGG
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → AAACTGCCTGTGGAAGATATTGTTAGCTCTCGTGAGTTTTCATTCGG
2590 2600 2610 2620 2630
AGGTAAGGAAGTGGATCAAGAACCTTCAGGCGAGGGCGTCACGAGAG
pDONR221_3HA_Toc159A.seq(1>4909) → AGGTAAGGAAGTGGATCAAGAACCTTCAGGCGAGGGCGTCACGAGAG
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → AGGTAAGGAAGTGGATCAAGAACCTTCAGGCGAGGGCGTCACGAGAG
2640 2650 2660 2670
TTGACGGaAGCGAAAGTGAGGAGGAAaCTGAGGAAATGATTTTCGGT
pDONR221_3HA_Toc159A.seq(1>4909) → TTGACGGAAGCGAAAGTGAGGAGGAAACTGAGGAAATGATTTTCGGT
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → TTGACGGAAGCGAAAGTGAGGAGGAAACTGAGGAAATGATTTTCGGT
Gatwy_3HA_Toc159A_cl1_P4.seq(1>895) → GG-AGCGAAAGTGAGGAGGAN-CTGAGGAAATGATTTTCGGT
2680 2690 2700 2710 2720
TCGTCTGAAGCTGCCAAACAATTCTTGGCTGAGCTTGAAAAAGCATC
pDONR221_3HA_Toc159A.seq(1>4909) → TCGTCTGAAGCTGCCAAACAATTCTTGGCTGAGCTTGAAAAAGCATC
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → TCGTCTGAAGCTGCCAAACAATTCTTGGCTGAGCTTGAAAAAGCATC
Gatwy_3HA_Toc159A_cl1_P4.seq(1>895) → TCGTCTGAAGCTGCCAAACAATTCTTGGCTGAGCTTGAAAAAGCATC
2730 2740 2750 2760 2770
TAGTGGAATCGAGGCTCATTCTGACGAAGCCAACATTTCCAACAACA
pDONR221_3HA_Toc159A.seq(1>4909) → TAGTGGAATCGAGGCTCATTCTGACGAAGCCAACATTTCCAACAACA
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → TAGTGGAATCGAGGCTCATTCTGACGAAGCCAACATTTCCAACAACA
Gatwy_3HA_Toc159A_cl1_P4.seq(1>895) → TAGTGGAATCGAGGCTCATTCTGACGAAGCCAACATTTCCAACAACA
2780 2790 2800 2810 2820
TGAGCGATAGAATTGACGGACAGATCGTGACTGACAGTGATGAAGAT
pDONR221_3HA_Toc159A.seq(1>4909) → TGAGCGATAGAATTGACGGACAGATCGTGACTGACAGTGATGAAGAT
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → TGAGCGATAGAATTGACGGACAGATCGTGACTGACAGTGATGAAGAT
Gatwy_3HA_Toc159A_cl1_P4.seq(1>895) → TGAGCGATAGAATTGACGGACAGATCGTGACTGACAGTGATGAAGAT
2830 2840 2850 2860
GTAGATACAGAAGATGAAGGTGAGGAAAAGATGTTCGATACAGCAGC
pDONR221_3HA_Toc159A.seq(1>4909) → GTAGATACAGAAGATGAAGGTGAGGAAAAGATGTTCGATACAGCAGC
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → GTAGATACAGAAGATGAAGGTGAGGAAAAGATGTTCGATACAGCAGC
Gatwy_3HA_Toc159A_cl1_P4.seq(1>895) → GTAGATACAGAAGATGAAGGTGAGGAAAAGATGTTCGATACAGCAGC
2870 2880 2890 2900 2910
ATTAGCAGCACTTTTAAAGGCAGCTACAGGGGGTGGGTCCTCAGAGG
pDONR221_3HA_Toc159A.seq(1>4909) → ATTAGCAGCACTTTTAAAGGCAGCTACAGGGGGTGGgTCcTCAGAGG
Gatwy_3HA_Toc159A_cl1_P3.seq(8>1059) → ATTAGCAGCACTTTTAAAGGCAGCTACAGGGGGTGGGTCCTCAGAGG
Gatwy_3HA_Toc159A_cl1_P4.seq(1>895) → ATTAGCAGCACTTTTAAAGGCAGCTACAGGGGGTGGGTCCTCAGAGG



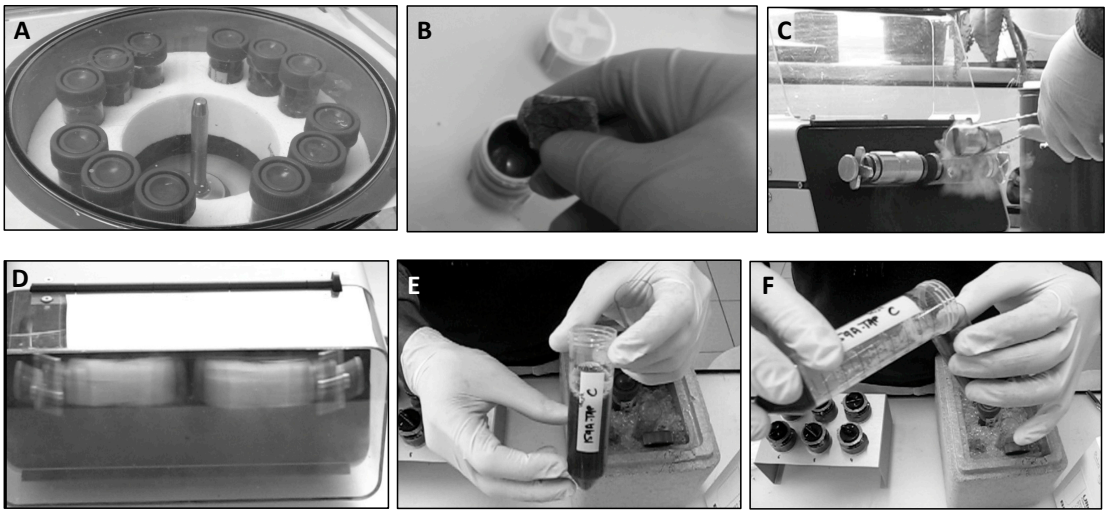
Appendix 1. The pEG101-3HA-TOC159A-YFP clone verified by sequencing.

Comparing the sequencing results(named Gatwy_) with the machine generated complete plasmid map pDONR221_3HA_TOC159A. The TOC159 sequence is from 753 bp to 3030bp. Several overlapping sequence results confirm the mutation free clone.



Appendix 2. Schematic diagram showing Toc159A with various modifications.

The diagram shows the sequence length, N-terminal and C-terminal modifications added to acidic domains of *TOC159* either wild type(wt) or the version with serine and threonine replaced by alanine (w/o ST).



Appendix 3 Formaldehyde crosslinking procedure in pictures.

A. Applying vacuum to the seedlings immersed in crosslinker solution

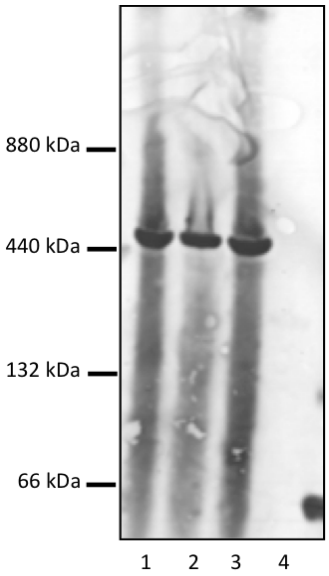
B. Adding the frozen plant samples into steel ball mill

C. Loading the steel ball mill on to Vibro-shaker

D. Vibro-shaker in function

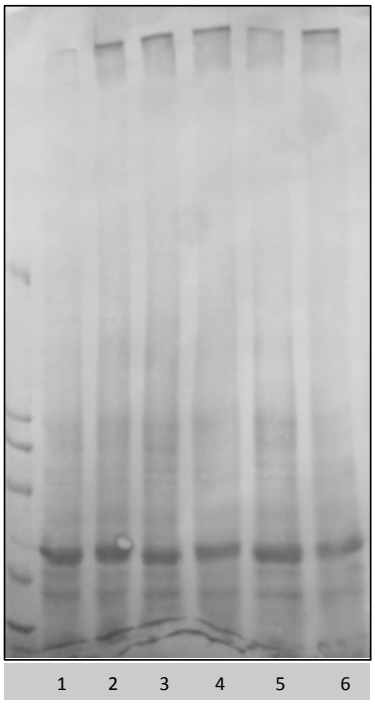
E. Resuspending the powdered sample in grinding buffer with protease inhibitor cocktail

F. Taking the crude extract supernatant for ultracentrifugation



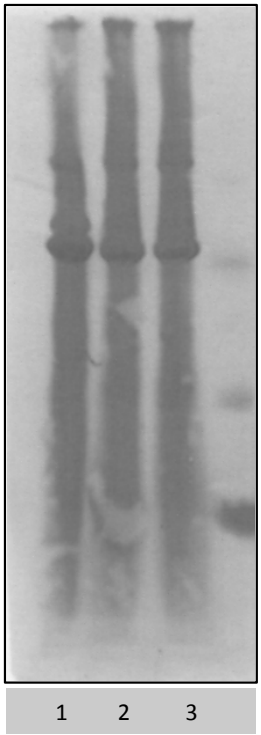
Appendix 4 Identification of Toc159A complex by BN-PAGE.

(linked to Fig.11) The BN-PAGE transferred to a nitrocellulose membrane and stained with coomassie blue to identify the lanes and compare protein load in each lane. Lane 1:Total protein from *TOC159A-TAP:WT,* lane2: Total protein from *TOC159A-TAP:WT* TAP tag specific proteins depleted using anti-CBP antibodies, lane 3: The same as in lane 1 but TAP-tag specific proteins blocked from entering gel by addition of anti-CBP antibodies, lane 4: recombinant 6xHis-Toc159A



Appendix 5 Formaldehyde crosslinked NTAP-Toc159 protein immuno blot Amidoblack staining

(linked to Fig. 12) Lane 1: No formaldehyde; lane 2: 0.5% formaldehyde; lane 3: 2% formaldehyde; lane 3: 4% formaldehyde; lane 5: 6% formaldehyde; lane 6: 10% formaldehyde.



Appendix 6. Formaldehyde crosslinked TOC159A-TAP:WT on BN-PAGE immuno blot coomassie stained

(linked to Fig 13) Lane 1: Non-crosslinked *TOC159A-TAP:WT;* lane2: 4% formaldehyde crosslinked *TOC159A-TAP:WT;* lane 3: 4% formaldehyde crosslinked *N-TAP:WT*

Interesting calculated features of Toc159A wt *(1-759 aa)*
Total number of negatively charged residues (Asp + Glu): 199
Total number of positively charged residues (Arg + Lys): 66

Atomic composition:

Carbon	C	3345
Hydrogen	H	5345
Nitrogen	N	881
Oxygen	O	1309
Sulfur	S	6

Formula: $C_{3345}H_{5345}N_{881}O_{1309}S_6$
Total number of atoms: 10886

Extinction coefficients: This protein does not contain any Trp residues. Experience shows that this could result in more than 10% error in the computed extinction coefficient.

Ext. coefficient 5960 Abs 0.1% (=1 g/l) 0.075

Instability index: The instability index (II) is computed to be 42.82
This classifies the protein as unstable.

Aliphatic index: 75.06
Grand average of hydropathicity (GRAVY): -0.576

Interesting calculated features of Toc159A w/o ST *(1-759 aa)*
Total number of negatively charged residues (Asp + Glu): 199
Total number of positively charged residues (Arg + Lys): 66
Atomic composition:

Carbon	C	3317
Hydrogen	H	5289
Nitrogen	N	881
Oxygen	O	1196
Sulfur	S	6

Formula: $C_{3317}H_{5289}N_{881}O_{1196}S_6$
Total number of atoms: 10689

Extinction coefficients:

This protein does not contain any Trp residues. Experience shows that this could result in more than 10% error in the computed extinction coefficient.

Ext. coefficient 5960 Abs 0.1% (=1 g/l) 0.078

Instability index:

The instability index (II) is computed to be 25.07
This classifies the protein as stable.

Aliphatic index: 89.95
Grand average of hydropathicity (GRAVY): -0.193

Source: ProtParam/ExPASY server (Wilkins et al., 1999)

7. References

Agne, B., & Kessler, F. (2010a). Modifications at the A-domain of the chloroplast import receptor Toc159. *Plant Signaling & Behavior*, *5*(11), 1513–1516. doi:10.4161/psb.5.11.13707

Agne, B., Andrès, C., Montandon, C., Christ, B., Ertan, A., Jung, F., et al. (2010b). The acidic A-domain of Arabidopsis TOC159 occurs as a hyperphosphorylated protein. *Plant Physiology*, *153*(3), 1016–1030. doi:10.1104/pp.110.158048

Agne, B., Infanger, S., Wang, F., Hofstetter, V., Rahim, G., Martin, M., et al. (2009). A toc159 import receptor mutant, defective in hydrolysis of GTP, supports preprotein import into chloroplasts. *The Journal of Biological Chemistry*, *284*(13), 8670–8679. doi:10.1074/jbc.M804235200

Andrès, C., Agne, B., & Kessler, F. (2010). The TOC complex: Preprotein gateway to the chloroplast. *Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research*, *1803*(6), 715–723. doi:10.1016/j.bbamcr.2010.03.004

Aronsson, H., & Jarvis, P. (2011). Dimerization of TOC receptor GTPases and its implementation for the control of protein import into chloroplasts. *Biochemical Journal*, *436*(2), e1–e2. doi:10.1042/BJ20110659

Aronsson, H., Combe, J., Patel, R., & Jarvis, P. (2006). In vivo assessment of the significance of phosphorylation of the Arabidopsis chloroplast protein import receptor, atToc33. *FEBS Letters*, *580*(2), 649–655. doi:10.1016/j.febslet.2005.12.055

Aronsson, H., Combe, J., Patel, R., Agne, B., Martin, M., Kessler, F., & Jarvis, P. (2010). Nucleotide binding and dimerization at the chloroplast pre-protein import receptor, atToc33, are not essential *in vivo* but do increase import efficiency. *The Plant Journal : for Cell and Molecular Biology*, *63*(2), 297–311. doi:10.1111/j.1365-313X.2010.04242.x

Asano, T., Yoshioka, Y., & Machida, Y. (2004). A defect in atToc159 of Arabidopsis thaliana causes severe defects in leaf development. *Genes & Genetic Systems*, *79*(4), 207–212.

Barsan, C., Zouine, M., Maza, E., Bian, W., Egea, I., Rossignol, 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 Physiology*, *160*(2), 708–725. doi:10.1104/pp.112.203679

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*(6766), 203–207. doi:10.1038/35003214

Bautista-Santos, A., & Zinker, S. (2014). The P1/P2 Protein Heterodimers Assemble to the Ribosomal Stalk at the Moment When the Ribosome Is Committed to Translation but Not to the Native 60S Ribosomal Subunit in Saccharomyces cerevisiae. *Biochemistry*, *53*(25), 4105–4112. doi:10.1021/bi500341w

Becker, T., Jelic, M., Vojta, A., Radunz, A., Soll, J., & Schleiff, E. (2004). Preprotein recognition by the Toc complex. *The EMBO Journal*, *23*(3), 520–530. doi:10.1038/sj.emboj.7600089

Bionda, T., TILLMANN, B., Simm, S., Beilstein, K., Ruprecht, M., & Schleiff, E. (2010). Chloroplast Import Signals: The Length Requirement for Translocation In Vitro and In Vivo. *Journal of Molecular Biology*, *402*(3), 510–523. doi:10.1016/j.jmb.2010.07.052

Bischof, S., Baerenfaller, K., Wildhaber, T., Troesch, R., Vidi, P. A., Roschitzki, B., et al. (2011). Plastid Proteome Assembly without Toc159: Photosynthetic Protein Import and Accumulation of N-Acetylated Plastid Precursor Proteins. *The Plant Cell Online*, *23*(11), 3911–3928. doi:10.1105/tpc.111.092882

Blobel, G. (1980). Regulation of intracellular protein traffic. *Harvey Lect*, *76*, 125–147.

Bölter, B., May, T., & Soll, J. (1998). A protein import receptor in pea chloroplasts, Toc86, is only a proteolytic fragment of a larger polypeptide. *FEBS Letters*, *441*(1), 59–62.

Brautigam, A., & Weber, A. P. M. (2009). Proteomic Analysis of the Proplastid Envelope Membrane Provides Novel Insights into Small Molecule and Protein Transport across Proplastid Membranes. *Molecular Plant*, *2*(6), 1247–1261. doi:10.1093/mp/ssp070

Bruce, B. D. (2001). The paradox of plastid transit peptides: conservation of function despite divergence in primary structure. *Biochimica Et Biophysica Acta*, *1541*(1-2), 2–21.

Chang, W., Soll, J., & Bölter, B. (2014). A new member of the psToc159 family contributes to distinct protein targeting pathways in pea chloroplasts. *Frontiers in Plant Science*, *5*, 239. doi:10.3389/fpls.2014.00239

Chavez, J. D., Liu, N. L., & Bruce, J. E. (2011). Quantification of protein-protein interactions with chemical cross-linking and mass spectrometry. *Journal of Proteome Research*, *10*(4), 1528–1537. doi:10.1021/pr100898e

Chen, K.-Y., & Li, H.-M. (2006). Precursor binding to an 880-kDa Toc complex as an early step during active import of protein into chloroplasts. *The Plant Journal*, *49*(1), 149–158. doi:10.1111/j.1365-313X.2006.02944.x

Chotewutmontri, P., Reddick, L. E., McWilliams, D. R., Campbell, I. M., & Bruce, B. D. (2012). Differential transit peptide recognition during preprotein binding and translocation into flowering plant plastids. *The Plant Cell Online*, *24*(7), 3040–3059. doi:10.1105/tpc.112.098327

Clough, S. J., & Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *The Plant Journal : for Cell and Molecular Biology*, *16*(6), 735–743.

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. *The Plant Journal*, *38*(1), 93–106. doi:10.1111/j.1365-313X.2004.02024.x

Dahlin, C., & Cline, K. (1991). Developmental Regulation of the Plastid Protein Import Apparatus. *The Plant Cell Online*, *3*(10), 1131–1140. doi:10.1105/tpc.3.10.1131

Danquah, A., de Zelicourt, A., Colcombet, J., & Hirt, H. (2014). The role of ABA and MAPK signaling pathways in plant abiotic stress responses. *Biotechnol Adv*, *32*(1), 40–52. doi:10.1016/j.biotechadv.2013.09.006

Dix, M. M., Simon, G. M., Wang, C., Okerberg, E., Patricelli, M. P., & Cravatt, B. F. (2012). Functional interplay between caspase cleavage and phosphorylation sculpts the apoptotic proteome. *Cell*, *150*(2), 426–440. doi:10.1016/j.cell.2012.05.040

Durek, P., Schmidt, R., Heazlewood, J. L., Jones, A., MacLean, D., Nagel, A., et al. (2010). PhosPhAt: the Arabidopsis thaliana phosphorylation site database. An update. *Nucleic Acids Research*, *38*(Database issue), D828–34. doi:10.1093/nar/gkp810

Dutta, S., Mohanty, S., & Tripathy, B. C. (2009). Role of Temperature Stress on Chloroplast Biogenesis and Protein Import in Pea. *Plant Physiology*, *150*(2), 1050–1061. doi:10.1104/pp.109.137265

Dutta, S., Teresinski, H. J., & Smith, M. D. (2014a). 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*(4), e95026. doi:10.1371/journal.pone.0095026.s004

Dutta, S., Teresinski, H. J., & Smith, M. D. (2014b). 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*(4), e95026. doi:10.1371/journal.pone.0095026

Ferro, M., Salvi, D., Brugiere, S., Miras, S., Kowalski, S., Louwagie, M., et al. (2003). Proteomics of the chloroplast envelope membranes from Arabidopsis thaliana. *Molecular & Cellular Proteomics : MCP*, *2*(5), 325–345. doi:10.1074/mcp.M300005-MCP200

Flores-Pérez, Ú., & Jarvis, P. (2013). Biochimica et Biophysica Acta. *BBA - Molecular Cell Research*, *1833*(2), 332–340. doi:10.1016/j.bbamcr.2012.03.019

Friso, G., Giacomelli, L., Ytterberg, A. J., Peltier, J. B., Rudella, A., Sun, Q., & Wijk, K. J. (2004). In-depth analysis of the thylakoid membrane proteome of Arabidopsis thaliana chloroplasts: new proteins, new functions, and a plastid proteome database. *The Plant Cell*, *16*(2), 478–499. doi:10.1105/tpc.017814

Fulgosi, H., & Soll, J. (2002). The chloroplast protein import receptors Toc34 and Toc159 are phosphorylated by distinct protein kinases. *The Journal of Biological Chemistry*, *277*(11), 8934–8940. doi:10.1074/jbc.M110679200

Gagat, P., Bodył, A., & Mackiewicz, P. (2013). How protein targeting to primary plastids via the endomembrane system could have evolved? A new hypothesis based on phylogenetic studies. *Biology Direct*, *8*, 18. doi:10.1186/1745-6150-8-18

Gibson, D. G., Smith, H. O., Hutchison, C. A. I., Venter, J. C., & Merryman, C. (2010). Chemical synthesis of the mouse mitochondrial genome. *Nature Methods*, *7*(11), 901–U905. doi:10.1038/nmeth.1515

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A. I., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), 343–U41. doi:10.1038/NMETH.1318

Gross, J., & Bhattacharya, D. (2009). Revaluating the evolution of the Toc and Tic protein translocons. *Trends in Plant Science*, *14*(1), 13–20. doi:10.1016/j.tplants.2008.10.003

Gutensohn, M., Schulz, B., Nicolay, P., & Flügge, U. I. (2000). Functional analysis of the two Arabidopsis homologues of Toc34, a component of the chloroplast protein import apparatus. *The Plant Journal : for Cell and Molecular Biology*, *23*(6), 771–783.

Heijne, von, G., & Nishikawa, K. (1991). Chloroplast transit peptides the perfect random coil? *FEBS Letters*, *278*(1), 1–3. doi:10.1016/j.tplants.2013.04.003

Hiltbrunner, A. (2001). Targeting of an abundant cytosolic form of the protein import receptor at Toc159 to the outer chloroplast membrane. *The Journal of Cell Biology*,

154(2), 309–316. doi:10.1083/jcb.200104022
Hiltbrunner, A., Bauer, J., Alvarez-Huerta, M., & Kessler, F. (2001). Protein translocon at the Arabidopsis outer chloroplast membrane. *Biochemistry and Cell Biology = Biochimie Et Biologie Cellulaire*, *79*(5), 629–635. doi:10.1139/bcb-79-5-629
Hirsch, S., Muckel, E., Heemeyer, F., Heijne, von, G., & Soll, J. (1994). A receptor component of the chloroplast protein translocation machinery. *Science (New York, N.Y.)*, *266*(5193), 1989–1992.
Hunter, T. (2007). The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Mol Cell*, *28*(5), 730–738. doi:10.1016/j.molcel.2007.11.019
Hust, B., & Gutensohn, M. (2006). Deletion of Core Components of the Plastid Protein Import Machinery Causes Differential Arrest of Embryo Development in Arabidopsis thaliana. *Plant Biology*, *8*(1), 18–30. doi:10.1055/s-2005-873044
Inaba, T. (2005). Arabidopsis Tic110 Is Essential for the Assembly and Function of the Protein Import Machinery of Plastids. *The Plant Cell Online*, *17*(5), 1482–1496. doi:10.1105/tpc.105.030700
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. *Molecular Plant*, *4*(2), 252–263. doi:10.1093/mp/ssq071
Inoue, H., Rounds, C., & Schnell, D. J. (2010). The Molecular Basis for Distinct Pathways for Protein Import into Arabidopsis Chloroplasts. *The Plant Cell Online*, *22*(6), 1947–1960. doi:10.1105/tpc.110.074328
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. *The Plant Journal : for Cell and Molecular Biology*, *34*(5), 661–669.
Ivanova, Y., Smith, M. D., Chen, K., & Schnell, D. J. (2004). Members of the Toc159 import receptor family represent distinct pathways for protein targeting to plastids. *Molecular Biology of the Cell*, *15*(7), 3379–3392. doi:10.1091/mbc.E03-12-0923
Jackson-Constan, D., & Keegstra, K. (2001). Arabidopsis genes encoding components of the chloroplastic protein import apparatus. *Plant Physiology*, *125*(4), 1567–1576.
Jarvis, P. (2008a). Targeting of nucleus-encoded proteins to chloroplasts in plants. *New Phytologist*, *179*(2), 257–285. doi:10.1111/j.1469-8137.2008.02452.x
Jarvis, P. (2008b). Targeting of nucleus-encoded proteins to chloroplasts in plants. *New Phytol*, *179*(2), 257–285.
Jarvis, P., & Lopez-Juez, E. (2013). Biogenesis and homeostasis of chloroplasts and other plastids. *Nature Reviews Molecular Cell Biology*, *14*(12), 787–802. doi:10.1038/nrm3702
Jarvis, P., Chen, L. J., Li, H., Peto, C. A., Fankhauser, C., & Chory, J. (1998). An Arabidopsis mutant defective in the plastid general protein import apparatus. *Science (New York, N.Y.)*, *282*(5386), 100–103.
Jelic, M., Soll, J., & Schleiff, E. (2003). Two Toc34 homologues with different properties. *Biochemistry*, *42*(19), 5906–5916. doi:10.1021/bi034001q
Jelic, M., Sveshnikova, N., Motzkus, M., Horth, P., Soll, J., & Schleiff, E. (2002). The chloroplast import receptor Toc34 functions as preprotein-regulated GTPase. *Biological Chemistry*, *383*(12), 1875–1883. doi:10.1515/BC.2002.211
Jensen, P. E., & Leister, D. (2014). Chloroplast evolution, structure and functions.

F1000prime Reports.
Jin, J. B., Kim, Y. A., Kim, S. J., Lee, S. H., Kim, D. H., Cheong, G.-W., & Hwang, I. (2001). A New Dynamin-Like Protein, ADL6, Is Involved in Trafficking from the trans-Golgi Network to the Central Vacuole in Arabidopsis.
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 Physiology*, *151*(3), 1339–1353. doi:10.1104/pp.109.145987
Kalanon, M., & McFadden, G. I. (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*(1), 95–112. doi:10.1534/genetics.107.085704
Kessler, F., & Schnell, D. J. (2002). A GTPase gate for protein import into chloroplasts. *Nature Structural Biology*, *9*(2), 81–83. doi:10.1038/nsb0202-81
Kessler, F., & Schnell, D. J. (2006). The Function and Diversity of Plastid Protein Import Pathways: A Multilane GTPase Highway into Plastids. *Traffic*, *7*(3), 248–257. doi:10.1111/j.1600-0854.2005.00382.x
Kessler, F., Blobel, G., Patel, H. A., & Schnell, D. J. (1994a). Identification of two GTP-binding proteins in the chloroplast protein import machinery. *Science (New York, N.Y.)*, *266*(5187), 1035–1039.
Kikuchi, S., Bedard, J., Hirano, M., Hirabayashi, Y., Oishi, M., Imai, M., et al. (2013). Uncovering the Protein Translocon at the Chloroplast Inner Envelope Membrane. *Science (New York, N.Y.)*, *339*(6119), 571–574. doi:10.1126/science.1229262
Kikuchi, S., Oishi, M., Hirabayashi, Y., Lee, D. W., Hwang, I., & Nakai, M. (2009). A 1-Megadalton Translocation Complex Containing Tic20 and Tic21 Mediates Chloroplast Protein Import at the Inner Envelope Membrane. *The Plant Cell Online*, *21*(6), 1781–1797. doi:10.1105/tpc.108.063552
Kleffmann, T., Russenberger, D., Zychlinski, von, A., Christopher, W., Sjölander, K., Gruissem, W., & Baginsky, S. (2004). The Arabidopsis thaliana Chloroplast Proteome Reveals Pathway Abundance and Novel Protein Functions. *Current Biology*, *14*(5), 354–362. doi:10.1016/j.cub.2004.02.039
Kleffmann, T., Zychlinski, von, A., Russenberger, D., Hirsch-Hoffmann, M., Gehrig, P., Gruissem, W., & Baginsky, S. (2007). Proteome dynamics during plastid differentiation in rice. *Plant Physiology*, *143*(2), 912–923. doi:10.1104/pp.106.090738
Klockenbusch, C., & Kast, J. (2010). Optimization of Formaldehyde Cross-Linking for Protein Interaction Analysis of Non-Tagged Integrin β1. *Journal of Biomedicine and Biotechnology*, *2010*, 1–14. doi:10.1155/2010/927585
Kovacheva, S., Bedard, J., Patel, R., & Dudley, P. (2005). In vivo studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. *The Plant …*.
Kovacs-Bogdan, E., Soll, J., & Bölter, B. (2010). Protein import into chloroplasts: the Tic complex and its regulation. *Biochimica Et Biophysica Acta*, *1803*(6), 740–747. doi:10.1016/j.bbamcr.2010.01.015
Kovács-Bogdán, E., Benz, J. P., Soll, J., & Bölter, B. (2011). Tic20 forms a channel independent of Tic110 in chloroplasts. *BMC Plant Biology*, *11*, 133. doi:10.1186/1471-2229-11-133
Kubis, S., Baldwin, A., Patel, R., Razzaq, A., Dupree, P., Lilley, K., et al. (2003). The

Arabidopsis ppi1 mutant is specifically defective in the expression, chloroplast import, and accumulation of photosynthetic proteins. *The Plant Cell*, *15*(8), 1859–1871.

Kubis, S., Patel, R., Combe, J., Bedard, J., Kovacheva, S., Lilley, K., et al. (2004). Functional Specialization amongst the Arabidopsis Toc159 Family of Chloroplast Protein Import Receptors. *The Plant Cell Online*, *16*(8), 2059–2077. doi:10.1105/tpc.104.023309

Lee, D. W., Jung, C., & Hwang, I. (2013). Biochimica et Biophysica Acta. *BBA - Molecular Cell Research*, *1833*(2), 245–252. doi:10.1016/j.bbamcr.2012.03.006

Lee, D. W., Lee, S., Oh, Y. J., & Hwang, I. (2009a). Multiple Sequence Motifs in the Rubisco Small Subunit Transit Peptide Independently Contribute to Toc159-Dependent Import of Proteins into Chloroplasts. *Plant Physiology*, *151*(1), 129–141. doi:10.1104/pp.109.140673

Lee, J., Kim, D. H., & Hwang, I. (2014). Specific targeting of proteins to outer envelope membranes of endosymbiotic organelles, chloroplasts, and mitochondria. *Frontiers in Plant Science*. doi:10.3389/fpls.2014.00173/abstract

Lee, J., Wang, F., & Schnell, D. J. (2009b). Toc Receptor Dimerization Participates in the Initiation of Membrane Translocation during Protein Import into Chloroplasts. *Journal of Biological Chemistry*, *284*(45), 31130–31141. doi:10.1074/jbc.M109.053751

Leister, D. (2003). Chloroplast research in the genomic age. *Trends Genet*, *19*(1), 47–56.

Li, H. M., & Chiu, C. C. (2010). Protein transport into chloroplasts. *Annual Review of Plant Biology*, *61*, 157–180. doi:10.1146/annurev-arplant-042809-112222

Li, H. M., & Teng, Y. S. (2013). Transit peptide design and plastid import regulation. *Trends in Plant Science*, *18*(7), 360–366. doi:10.1016/j.tplants.2013.04.003

Ling, Q., Huang, W., Baldwin, A., & Jarvis, P. (2012a). Chloroplast Biogenesis Is Regulated by Direct Action of the Ubiquitin-Proteasome System. *Science (New York, N.Y.)*, *338*(6107), 655–659. doi:10.1126/science.1225053

Markwell, M. A., & Fox, C. F. (1980). Protein-protein interactions within paramyxoviruses identified by native disulfide bonding or reversible chemical cross-linking. *Journal of Virology*, *33*(1), 152–166.

Olsen, L. J., & Keegstra, K. (1992). The binding of precursor proteins to chloroplasts requires nucleoside triphosphates in the intermembrane space. *The Journal of Biological Chemistry*, *267*(1), 433–439.

Oreb, M., Hofle, A., Koenig, P., Sommer, M. S., Sinning, I., Wang, F., et al. (2011). Substrate binding disrupts dimerization and induces nucleotide exchange of the chloroplast GTPase Toc33. *Biochemical Journal*, *436*(2), 313–319. doi:10.1042/BJ20110246

Oreb, M., Hofle, A., Mirus, O., & Schleiff, E. (2008). Phosphorylation regulates the assembly of chloroplast import machinery. *Journal of Experimental Botany*, *59*(9), 2309–2316. doi:10.1093/jxb/ern095

Oreb, M., Zoryan, M., Vojta, A., Maier, U. G., Eichacker, L. A., & Schleiff, E. (2007). Phospho-mimicry mutant of atToc33 affects early development of Arabidopsis thaliana. *FEBS Letters*, *581*(30), 5945–5951. doi:10.1016/j.febslet.2007.11.071

Perry, S. E., & Keegstra, K. (1994). Envelope membrane proteins that interact with chloroplastic precursor proteins. *The Plant Cell*, *6*(1), 93–105. doi:10.1105/tpc.6.1.93

Phizicky, E. M., & Fields, S. (1995). Protein-protein interactions: methods for detection and analysis. *Microbiological Reviews*, *59*(1), 94–123.

Primavesi, L. F., Wu, H., Mudd, E. A., Day, A., & Jones, H. D. (2008). Visualisation of plastids in endosperm, pollen and roots of transgenic wheat expressing modified GFP fused to transit peptides from wheat SSU RubisCO, rice FtsZ and maize ferredoxin III proteins. *Transgenic Res*, *17*(4), 529–543. doi:10.1007/s11248-007-9126-7

Rahim, G., Bischof, S., Kessler, F., & Agne, B. (2009). In vivo interaction between atToc33 and atToc159 GTP-binding domains demonstrated in a plant split-ubiquitin system. *Journal of Experimental Botany*, *60*(1), 257–267. doi:10.1093/jxb/ern283

Reumann, S., Inoue, K., & Keegstra, K. (2005). Evolution of the general protein import pathway of plastids (review). *Mol Membr Biol*, *22*(1-2), 73–86.

Richardson, L. G., Jelokhani-Niaraki, M., & Smith, M. D. (2009). The acidic domains of the Toc159 chloroplast preprotein receptor family are intrinsically disordered protein domains. *BMC Biochemistry*, *10*(1), 35. doi:10.1186/1471-2091-10-35

Schleiff, E., & Soll, J. (2005). Membrane protein insertion: mixing eukaryotic and prokaryotic concepts. *EMBO Reports*, *6*(11), 1023–1027. doi:10.1038/sj.embor.7400563

Schnell, D. J., Kessler, F., & Blobel, G. (1994). Isolation of components of the chloroplast protein import machinery. *Science (New York, N.Y.)*, *266*(5187), 1007–1012.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., & Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols*, *1*(6), 2856–2860. doi:10.1038/nprot.2006.468

Shevchenko, A., Wilm, M., Vorm, O., & Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Analytical Chemistry*, *68*(5), 850–858.

Shi, L. X., & Theg, S. M. (2013). Energetic cost of protein import across the envelope membranes of chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(3), 930–935. doi:10.1073/pnas.1115886110

Shi, X., Zhang, H., & Lin, S. (2013). Tandem repeats, high copy number and remarkable diel expression rhythm of form II RuBisCO in Prorocentrum donghaiense (Dinophyceae). *PLoS ONE*, *8*(8), e71232. doi:10.1371/journal.pone.0071232

Smith, D. R., & Lee, R. W. (2014). A Plastid without a Genome: Evidence from the Nonphotosynthetic Green Algal Genus Polytomella. *Plant Physiology*.

Smith, M. D. (2002). The targeting of the atToc159 preprotein receptor to the chloroplast outer membrane is mediated by its GTPase domain and is regulated by GTP. *The Journal of Cell Biology*, *159*(5), 833–843. doi:10.1083/jcb.200208017

Smith, M. D. (2004). atToc159 is a selective transit peptide receptor for the import of nucleus-encoded chloroplast proteins. *The Journal of Cell Biology*, *165*(3), 323–334. doi:10.1083/jcb.200311074

Smith, M. D., Schnell, D. J., Fitzpatrick, L., & Keegstra, K. (2003). In vitro analysis of chloroplast protein import. *Current Protocols in Cell Biology / Editorial Board, Juan S. Bonifacino ... [Et Al.]*, *Chapter 11*, Unit11.16. doi:10.1002/0471143030.cb1116s17

Sun, C. W., Chen, L. J., Lin, L. C., & Li, H. M. (2001). Leaf-specific upregulation of chloroplast translocon genes by a CCT motif-containing protein, CIA 2. *The Plant Cell*, *13*(9), 2053–2061.

Sun, C. W., Huang, Y. C., & Chang, H. Y. (2009). CIA2 coordinately up-regulates protein import and synthesis in leaf chloroplasts. *Plant Physiology*, *150*(2), 879–888. doi:10.1104/pp.109.137240

Sun, Y. J., Forouhar, F., Li Hm, H. M., Tu, S. L., Yeh, Y. H., Kao, S., et al. (2002). Crystal structure of pea Toc34, a novel GTPase of the chloroplast protein translocon. *Nature Structural Biology*, *9*(2), 95–100. doi:10.1038/nsb744

Sutherland, B. W., Toews, J., & Kast, J. (2008). Utility of formaldehyde cross-linking and mass spectrometry in the study of protein–protein interactions. *Journal of Mass Spectrometry*, *43*(6), 699–715. doi:10.1002/jms.1415

Sveshnikova, N., Grimm, R., Soll, J., & Schleiff, E. (2000). Topology Studies of the Chloroplast Protein Import Channel Toc75. *Biological Chemistry*, *381*(8), 687–693. doi:10.1515/BC.2000.089

Tchórzewski, M. (2002). The acidic ribosomal P proteins. *The International Journal of Biochemistry & Cell Biology*, *34*(8), 911–915. doi:10.1016/S1357-2725(02)00012-2

Teng, Y.-S., Chan, P.-T., & Li, H.-M. (2012). Differential age-dependent import regulation by signal peptides. *PLoS Biology*, *10*(10), e1001416. doi:10.1371/journal.pbio.1001416

Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M., & Tommassen, J. (2003). Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science (New York, N.Y.)*, *299*(5604), 262–265. doi:10.1126/science.1078973

Waegemann, K., & Soil, J. (1991). Characterization of the protein import apparatus in isolated outer envelopes of chloroplasts. *The Plant Journal*, *1*(2), 149–158. doi:10.1111/j.1365-313X.1991.00149.x

Wallas, T. R., Smith, M. D., Sanchez-Nieto, S., & Schnell, D. J. (2003). The roles of toc34 and toc75 in targeting the toc159 preprotein receptor to chloroplasts. *The Journal of Biological Chemistry*, *278*(45), 44289–44297. doi:10.1074/jbc.M307873200

Wan, J., Blakeley, S. D., Dennis, D. T., & Ko, K. (1996). Transit peptides play a major role in the preferential import of proteins into leucoplasts and chloroplasts. *The Journal of Biological Chemistry*, *271*(49), 31227–31233.

Wang, P., Xue, L., Batelli, G., Lee, S., Hou, Y. J., Van Oosten, M. J., et al. (2013). Quantitative phosphoproteomics identifies SnRK2 protein kinase substrates and reveals the effectors of abscisic acid action. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(27), 11205–11210. doi:10.1073/pnas.1308974110

Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., & Hochstrasser, D. F. (1999). Protein identification and analysis tools in the ExPASy server. *Methods in Molecular Biology (Clifton, N.J.)*, *112*, 531–552.

Yan, J., Campbell, J. H., Glick, B. R., Smith, M. D., & Liang, Y. (2014). Molecular characterization and expression analysis of chloroplast protein import components in tomato (Solanum lycopersicum). *PLoS ONE*, *9*(4), e95088. doi:10.1371/journal.pone.0095088

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 Physiology*, *127*(1), 90–96.

Zhang, X. P., & Glaser, E. (2002). Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone. *Trends in Plant Science*, *7*(1), 14–21.

Zhong, R., Thompson, J., Ottesen, E., & Lamppa, G. K. (2010). A forward genetic screen to explore chloroplast protein import *in vivo* identifies Moco sulfurase, pivotal for ABA and IAA biosynthesis and purine turnover. *The Plant Journal : for Cell and Molecular Biology*, *63*(1), 44–59. doi:10.1111/j.1365-313X.2010.04220.x

Zhu, L., Li, M., Wei, L., Liu, X., Yin, J., & Gao, Y. (2014). Fast fixing and comprehensive identification to help improve real-time ligands discovery based on formaldehyde crosslinking, immunoprecipitation and SDS-PAGE separation. *Proteome Science*, *12*(1), 1–8. doi:10.1186/1477-5956-12-6

Acknowledgments

First of all, I would like to thank Prof. Felix Kessler for giving me the opportunity to perform my Ph.D. thesis in his group at the Laboratory of Plant Physiology, University of Neuchâtel. I highly appreciate his communicative enthusiasm and his numerous motivating ideas. I am deeply grateful for his support and encouragement throughout my thesis work, as well as for sharing his knowledge and experience on several curricular and extracurricular topics.

I would like to thank Dr. Birgit Agne for all the help she offered when I landed on this project. I am immensely grateful to her continued input during our online discussion and collaboration on the proteomics front even after she moved to Halle.

I appreciate very much our collaboration with Prof. Sacha Baginsky for proteomics solutions.

I would also like to acknowledge Prof. Jean-Marc Neuhaus, Prof. Samuel Zeeman and Prof. Dr. Paul Jarvis for accepting to serve in my thesis committee.

I am much obliged to Dr. Emilie Demarsy for her help in the lab and in shaping the structure of this thesis. I also appreciate her contribution towards our joint publication.

I would like to thank Veronique Douet for offering her technical skills in BN-PAGE and protein import, microscopy and for taking care of the supply of necessary reagents in the lab. I also extend my thanks to Dr. Charles Andreas and Dr. Céline Besagni for offering help in microscopy and in other discussions in the lab.

In general many thanks to all the past and present members of LPV, at University of Neuchâtel. In particular I am deeply grateful to Cyrille Montandon for our long discussion in the lab on trouble shooting, sharing his technical experience and knowledge in the lab and outside. I am also grateful to Meryl Martin and Shanmugabalaji for their support in the beginning of my thesis..

Particularly, I am very much indebted to my parents who made all this possible; I am grateful to my parents and my sister for their unconditional love and support.

Last but not least, for a person like me who is living far from home it is all the love and support of friends that helped me through all the way from the beginning of my start in Switzerland. My deep love and gratitude goes to them.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 17 September 2014
doi: 10.3389/fpls.2014.00483

Border control: selectivity of chloroplast protein import and regulation at the TOC-complex

Emilie Demarsy†, Ashok M. Lakshmanan† and Felix Kessler*

Laboratory of Plant Physiology, University of Neuchâtel, Neuchâtel, Switzerland

Edited by:
Kentaro Inoue, University of California at Davis, USA

Reviewed by:
Matthew D. Smith, Wilfrid Laurier University, Canada
Takehito Inaba, University of Miyazaki, Japan

***Correspondence:**
Felix Kessler, Laboratory of Plant Physiology, Université de Neuchâtel, UniMail, Rue Emile Argand 11, 2000 Neuchâtel, Switzerland
e-mail: felix.kessler@unine.ch

†Emilie Demarsy and Ashok M. Lakshmanan have contributed equally to this work.

Plants have evolved complex and sophisticated molecular mechanisms to regulate their development and adapt to their surrounding environment. Particularly the development of their specific organelles, chloroplasts and other plastid-types, is finely tuned in accordance with the metabolic needs of the cell. The normal development and functioning of plastids require import of particular subsets of nuclear encoded proteins. Most preproteins contain a cleavable sequence at their N terminal (transit peptide) serving as a signal for targeting to the organelle and recognition by the translocation machinery TOC–TIC (translocon of outer membrane complex–translocon of inner membrane complex) spanning the dual membrane envelope. The plastid proteome needs constant remodeling in response to developmental and environmental factors. Therefore selective regulation of preprotein import plays a crucial role in plant development. In this review we describe the diversity of transit peptides and TOC receptor complexes, and summarize the current knowledge and potential directions for future research concerning regulation of the different Toc isoforms.

Keywords: plastids, protein import, TOC complex, preproteins, post-translational modifications

INTRODUCTION

Eukaryotic cells are composed of multiple compartments that acquire specialized sets of proteins for function. The vast majority of proteins are encoded by the nuclear genome. After synthesis in the cytosol accurate protein sorting and export toward their destination organelles rely on intrinsic topogenic sequences (Blobel, 1980). Initially, correct recognition of a preprotein requires specific receptors at the surface of the organelle. This crucial step of intracellular trafficking control can be viewed as a key–lock type mechanism.

Plant chloroplasts import impressive quantities as well as an enormous diversity of proteins from the cytosol. Large scale proteome studies indicate that 2000–4000 different proteins follow the chloroplast route (Ferro et al., 2003; Leister, 2003; Friso et al., 2004; Kleffmann et al., 2004). In the cytosol, chloroplast proteins are generally synthesized as preproteins with a N-terminal targeting sequence that is cleaved to produce the mature chloroplast protein upon import. This N-terminal targeting sequence, named transit peptide in the context of chloroplast protein import, faithfully guides the preprotein to the chloroplast surface where it engages the import machinery. In the following, the preprotein is translocated across the dual envelope membranes into the stroma. The transit sequence is cleaved upon arrival in the stroma yielding the mature form of the protein followed by folding in the stroma, targeting to the inner membrane via the conservative sorting pathway, or transport to the thylakoid membrane system. The recognition and translocation of the preprotein at the plastid envelope is provided by the TOC–TIC (translocon of outer membrane complex–translocon of inner membrane complex (TOC–TIC) import machinery. In pea, the core TOC complex consists of an assembly of the two GTP dependent receptors Toc34 and Toc159 together with the β-barrel protein conducting channel Toc75 (Hirsch et al., 1994; Kessler et al., 1994; Perry and Keegstra, 1994; Schnell et al., 1994; Becker et al., 2004). Upon engagement of the preprotein, the TOC complex associates with the TIC complex to form a continuous channel through the plastid envelope. The protein conducting channel at the TIC complex has been suggested to be made up of Tic110 or Tic20, or yet a combination of the two. Recently, however, it has been suggested that four core components form a 1MDa TIC channel [Tic20, Tic214 formerly known as YCF1, Tic56, and Tic100; (Kikuchi et al., 2013)]. Protein synthesis and targeting involve a large variety of cellular activities that are energy-requiring. Solely translocation of a single preprotein across the chloroplast envelope through the TOC–TIC machinery requires the hydrolysis of 650 ATP molecules on average, representing about 0.6% of the total light-saturated energy output of the organelles (Shi and Theg, 2013). Therefore a tight control of TOC–TIC mediated import activity is required to respect the cellular energy budget allocated to protein import.

Plants originate from a primary endosymbiotic event involving a photosynthetic cyanobacterium captured by a eukaryotic cell. The evolution of plants toward complex and multicellular organisms has been accompanied by the diversification of interconvertible plastid types displaying distinct and highly specialized biochemical and physiological functions (Jarvis and Lopez-Juez, 2013). For instance the most prominent plastid type, the chloroplast, develop from proplastid, or partially differentiated, non-photosynthetic etioplast, and can also differentiate into other non-photosynthetic plastid types such as chromoplast or elaioplast. Each plastid type requires the import of different subsets of proteins (Kleffmann et al., 2007; Brautigam and

Weber, 2009; Barsan et al., 2012). Several strategies have evolved coordinately to ensure the selective import of plastid proteins. Together with the defined regulation of preprotein availability at the transcriptional levels, evolution also triggered diversification and increased complexity of both preprotein transit sequences (von Heijne and Nishikawa, 1991; Li and Teng, 2013) and composition of the import machinery (Reumann et al., 2005; Kalanon and McFadden, 2008; Gross and Bhattacharya, 2009; Shi and Theg, 2013). Evidence for the existence of different isoforms of TOC complex components has now been reported for several higher plant species including *Arabidopsis*, pea, and tomato (Jackson-Constan and Keegstra, 2001; Chang et al., 2014; Yan et al., 2014). Each isoform is thought to preferentially import a specific subset of client preproteins that may be the result of differential binding affinity (Jelic et al., 2003; Smith et al., 2004; Inoue et al., 2010; Dutta et al., 2014). Therefore, the relative abundance of Toc isoforms may reflect the protein composition of a given plastid type and be a key marker of plastid identity (Ling et al., 2012).

On top of that, plants are sessile organisms and need to adapt to ever-changing environmental conditions. Dynamic regulation of TOC complex composition may occur at the posttranslational level and represent a key regulatory mechanism contributing to the change in protein composition. By consequence this allows rapid modulation of plastid metabolism to ensure and drive plant development and acclimation. Thus, the relative abundance of Toc receptor may not only be a marker of plastid type but also of plastid state (Agne et al., 2010; Ling et al., 2012).

The molecular mechanisms underlying the process of protein translocation have been reviewed extensively (Jarvis, 2008; Andres et al., 2010; Li and Chiu, 2010). Here, we present the current knowledge with regard to the selectivity and the regulation of the preprotein import process at the level of the TOC complex.

PREPROTEIN IMPORT IN PLASTID IS REGULATED BY DEVELOPMENTAL AND ENVIRONMENTAL FACTORS

Years before the identification of any of the components of the chloroplast protein import machinery (Dahlin and Cline, 1991) proposed that import activity is correlated with protein demands during plastid development. They observed a high import activity in non-photosynthetic proplastids, which gradually decreased as plastids matured. This phenomenon was observed for etioplast as well as chloroplast development. Interestingly, when dark-grown plants were shifted from dark to light the import activity of etioplasts was activated to accommodate the set of preproteins required for chloroplast differentiation (Dahlin and Cline, 1991). This seminal study focused on a few substrates and, given the experimental limitations at the time, was unable to provide a complete picture of plastid protein import dynamics. Recently, this topic was reinvestigated using a larger number of chloroplasts precursors proteins (Teng et al., 2012). This study confirmed that preprotein specificity is modulated in synchrony with chloroplast developmental stages. Interestingly, this study demonstrated that the earlier results by Dahlin and Cline (1991) cannot be extended to all import substrates. Rather, Teng et al. (2012) refined the model and classified the substrates according to their importability in chloroplasts at different developmental stages and consequently defined three age-selective classes: substrates that are imported more efficiently in young chloroplasts (group I), in older chloroplasts (group III), whereas group II represents substrates that are imported similarly in developing and mature chloroplasts. Thus, it appears that regulation of chloroplast preprotein import is part of a differential age-specific regulatory network.

In vitro import experiments using different isolated plastid types as well as the visualization of protein targeting using transgenic lines expressing transit peptides fused to GFP support the notion that import selectivity is regulated in a tissue specific manner (Wan et al., 1996; Primavesi et al., 2008; Yan et al., 2014). Finally temperature stress (cold and heat) on intact pea leaves and isolated chloroplasts was found to reduce import of the small subunit of RubisCO preprotein (pSSU; Dutta et al., 2009).

In summary, these results demonstrate that both plastid import activity and selectivity are modulated in accordance with plastid type, developmental stage, and environmental condition. For this purpose plants have evolved a complex set of preprotein import components with specialized features and regulatory mechanisms (Jarvis et al., 1998; Kubis et al., 2004).

PREPROTEIN SELECTIVITY AT THE CHLOROPLAST IMPORT MACHINERY

OVERVIEW OF THE TOC–TIC MACHINERY

The TOC-TIC pathway (translocon of outer membrane complex-translocon of inner membrane complex) is the major protein import pathway in higher plants (Bauer et al., 2000; Asano et al., 2004; Kovacheva et al., 2005). Most of the proteins with cleavable transit peptides that are targeted to the stroma, thylakoid membranes and lumen follow this route, that is therefore vital for plastid biogenesis (Kessler and Schnell, 2006; Bischof et al., 2011; Dutta et al., 2014). The native TOC–TIC complex in pea and *Arabidopsis* has been found to include two GTPase-receptors Toc159 and Toc34, a channel protein Toc75 and at least three additional regulatory Toc proteins Toc64, Toc22, and Toc12 (Andres et al., 2010). At the inner membrane at least 11 different proteins have been reported to be involved in the import process (Kovacs-Bogdan et al., 2010; Kikuchi et al., 2013). Electrophysiological experiments suggested that Tic110 and Tic20 could function as channels facilitating the translocation of preproteins across the inner membrane (Kikuchi et al., 2009, 2013; Kovacs-Bogdan et al., 2011). These two channels are thought to function independently and in different complexes (Kikuchi et al., 2009, 2013; Kovacs-Bogdan et al., 2011). This is supported by the finding that Tic110 interacts with preproteins and TOC complexes (Schnell et al., 1994; Inaba et al., 2005) but not with Tic20 (Kikuchi et al., 2009). Tic110 is a protein of eukaryotic origin present in various plastid-containing organisms (Shi and Theg, 2013). Its function is indispensable for plant viability and chloroplast biogenesis (Inaba et al., 2005). Based on these data it was proposed that Tic110 has an essential role in chloroplast protein import. Recently, composition of the Tic20 complex in *Arabidopsis* has been investigated using Blue Native PAGE and mass spectrometric analyses. The results suggested that Tic20 associates with Tic56, Tic100, and Tic214 (Kikuchi et al., 2013). Although Tic20 is of prokaryotic origin and is well conserved

among the plant kingdom, Tic56, Tic100, and Tic214 appear to have specifically evolved in a limited number of higher plant species only (Kikuchi et al., 2013). Tic214, also known as YCF1, is absent from the genome of some Poacae species (Jensen and Leister, 2014; Smith and Lee, 2014), thus the role of TIC20 complex as the general inner chloroplast membrane translocon in higher plants is questionable. Nevertheless the albino, seedling lethal phenotype of null mutants of each of the TIC20 complex subunits underscores their functional importance at least in *Arabidopsis*. In conclusion the exact contribution of TIC110 and TIC20 complexes in chloroplast protein import is still under debate.

At the evolutionary level, a view of growing complexity of the composition of TOC machinery is emerging (Shi and Theg, 2013). Starting with one channel protein at the outer envelope in cyanobacteria, the outer envelope protein import complex has evolved into a GTP-regulated multi-protein complex in higher plants (Olsen and Keegstra, 1992; Schnell et al., 1994; Hiltbrunner et al., 2001a; Kessler and Schnell, 2002; Voulhoux et al., 2003). The Toc receptors can form homo- and heterodimers in a dynamic way regulated by preprotein binding and GTP binding/hydrolysis activity (Smith et al., 2002; Sun et al., 2002; Wallas et al., 2003; Lee et al., 2009b; Rahim et al., 2009; Oreb et al., 2011). Although GTP binding and GTPase activity seem dispensable (expression of GTPase/dimerization-defective Toc159 and Toc33 complement the corresponding knock out mutants), it has been shown that they are required for full preprotein import efficiency *in vitro* (Agne et al., 2009; Aronsson et al., 2010; Aronsson and Jarvis, 2011). In most higher plants the Toc75 channel is encoded by a single gene (Inoue and Keegstra, 2003), but normally more than one homolog for the plastid specific GTPase families Toc159 and Toc34 exists, and thus there is the possibility of making various combinations of TOC complexes (Hiltbrunner et al., 2001a; Chang et al., 2014; Yan et al., 2014). The evolution of a translocation route depending on GTP-binding as well as other accessory proteins may be seen as the key to the developmental stage specific regulation of protein import in higher plants (Schleiff and Soll, 2005; Gagat et al., 2013).

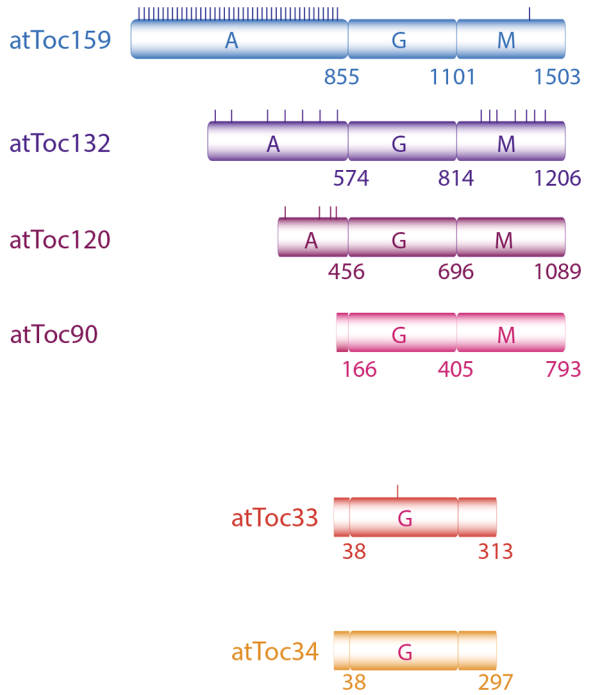


FIGURE 1 | The translocon of outer membrane complex includes two GTPase-receptors Toc159 and Toc34. Toc159 consists of a GTPase domain (G) flanked by a C-terminal membrane anchoring domain (M) and an acidic N-terminal region (A). In *Arabidopsis* four Toc159 and two Toc34 isoforms have been identified. Toc159 homologs differ primarily in their A-domain sequences and lengths. Experimentally identified phospho-serine and -threonine residues (PhosphAT 4.0) are schematically indicated by short vertical lines.

DIVERSITY AND FUNCTIONAL SPECIFICITIES OF TOC GTPase RECEPTORS

Members of Toc159 family are characterized by three distinct domains: M- (membrane anchoring) domain, G- (GTP-binding) domain, and a highly acidic, intrinsically disordered A-domain (**Figure 1**). There are four homologs in *Arabidopsis thaliana*: atToc159, −132, −120, and −90. While they share high similarity in their G- and M-domains, they largely differ in length and sequence at their A-domains (Jackson-Constan and Keegstra, 2001; Hiltbrunner et al., 2001a). Toc34 proteins are smaller, membrane-anchored GTPases. In pea, only one member has been detected so far while two isoforms of Toc34 (atToc34 and atToc33) have been identified in *Arabidopsis* (Jarvis et al., 1998; Gutensohn et al., 2000).

Genetics and biochemical studies have supported the idea that various combinations of the different Toc GTPase isoforms lead to a diversity of complexes displaying differential selectivity for preprotein recognition and translocation (Kubis et al., 2003, 2004; Constan et al., 2004; Ivanova et al., 2004). Co-immunoprecipitation experiments performed by Ivanova and collaborators demonstrated that atToc159 preferentially associates with atToc33, while atToc120, and/or atToc132 preferentially form a complex together with atToc34 (Ivanova et al., 2004). Interestingly, the *toc34 (ppi3)* knock out mutant has no visible defect, while the *toc33 (ppi1)* mutant displays a pale green phenotype with a chloroplast biogenesis defect similar (although much less severe) than the *toc159* mutant phenotype (*ppi2*), supporting the proposition that these latter two receptor isoforms function in the same complex and preprotein import pathway (Jarvis et al., 1998; Bauer et al., 2000; Kubis et al., 2003, 2004; Constan et al., 2004).

Several lines of evidence indicate a potential functional overlap of the two Toc34 members: the strong sequence similarity: 65% (Jarvis et al., 1998); the fact that a minor fraction of atToc33 was co-immunoprecipated with Toc120/132, and atToc34 was detected with atToc159 (Ivanova et al., 2004); the embryo lethal phenotype of *toc33/toc34* double mutants and, most importantly, the ability of atToc34 to complement *ppi1* phenotype (Jarvis et al., 1998; Kubis et al., 2003; Constan et al., 2004). Transgenic complementation studies also indicated the potential functional overlap of atToc120 and atToc132 (Ivanova et al., 2004; Kubis et al., 2004) and, to a limited extent, for atToc159 and atToc90 (Infanger et al.,

2011), however, no functional overlap exists between these two subgroups [atToc120/132 vs. atToc159/atToc90 (Ivanova et al., 2004; Kubis et al., 2004)]. While the two Toc34 homologs are mutually exchangeable, the same is only partially true for the Toc159 homologs, suggesting that preprotein selectivity of TOC complexes is mostly conferred by the identity of the Toc159 isoforms.

The classification of the client proteins of each isoform has been attempted. Because of the albino phenotype of *ppi2*, it has been proposed that Toc159 primarily facilitates the import of photosynthesis-associated preproteins. On the other hand, Toc132, or Toc120 being present predominantly in roots could facilitate that of constitutive (housekeeping) preproteins (Kubis et al., 2003, 2004; Ivanova et al., 2004; Smith et al., 2004; Inoue et al., 2010). *In vitro* import assays using a selection of substrates support this model (Smith et al., 2004; Inoue et al., 2010). However, the albino phenotype of the *ppi2* mutant was shown to result not only from a defect in the import of a set of chloroplast proteins, but also from the transcriptional downregulation of a specific set of nuclear genes associated with photosynthesis (Bauer et al., 2000; Kakizaki et al., 2009). This effect is commonly referred to as retrograde signaling, and pleiotropically affects albino and pale green mutants across the board. The interference of retrograde signaling with preprotein import in *ppi* mutants has blurred the identification of the specific substrates of each of the receptor isoforms. Comparative analysis of *ppi2* mutant proteome and transcriptome demonstrated that certain photosynthesis-associated proteins accumulated normally in plastids even in the absence of atToc159, whereas accumulation of some house-keeping proteins were strongly diminished despite their mRNA expression levels being similar to the wild type (Bischof et al., 2011). Furthermore, the results of a yeast two hybrid screen used to identify the preferred Toc receptor of a variety of preproteins supported to the finding of (Bischof et al., 2011; Dutta et al., 2014). Together these studies affirmed that Toc GTPases, especially the Toc159 homologs, confer specificity to plastid preprotein import. However, specificity is not likely to be based on the photosynthetic or housekeeping nature of a preprotein. This is a move away from the overly simplistic paradigm of "photosynthesis-associated" and "house-keeping" specificities toward a more differentiated model that reflects complex and varying plastid preprotein requirements during development and under environmental influence. Therefore, Toc client protein classification will need to be rethought along these lines. One hypothesis is that the combination of preprotein specificities of plastid resident Toc receptors reflects the tissue or cell specific preprotein accumulation patterns that are specific to a particular plastid type.

As mentioned above Toc159 homologs diverge the most at their A-domains, suggesting a key role in their functional specialization. In domain swapping experiments, Inoue et al. (2010) replaced the A-domain of atToc132 by that of atToc159. Expression of this construct partially restored chlorophyll accumulation in the *toc159* null mutant (*ppi2*), while no complementation was observed using a construct encoding atToc132 without an A-domain. These data elegantly demonstrated that the functional specialization relies at least partially on intrinsic properties of the A-domain (Inoue et al., 2010). In agreement with this, it was observed that removal of the A-domains of atToc159 and atToc132 reduced the binding selectivity of these isoforms (Smith et al., 2004; Inoue et al., 2010; Dutta et al., 2014). Apparently, the A-domain does not directly interact with preproteins but may act as a filter enhancing the affinity for subsets of proteins and reducing the affinity for others (Dutta et al., 2014). Preprotein binding to Toc159 has been shown earlier to occur at the G-domain (Smith et al., 2004). Thus it seems likely that the A-domain influences the G-domain by, for instance, positively, or negatively modulating access of a preprotein according to its nature. Finally, the lack of complementation of *ppi2* by atToc132 lacking an A-domain (Inoue et al., 2010) as well as the recent work of Smith et al. (2004) using a yeast two hybrid system to study the preprotein-Toc159 receptor isoforms affinity (Dutta et al., 2014) indicate that a degree of specificity is conferred by the G-domain itself.

DIVERSITY AND COMPLEXITY OF THE TRANSIT PEPTIDES

Inherently, recognizable specificity features would need to be encoded in the plastid transit peptides. One general consideration regarding the transit peptides is that no consensus can be defined, even when considering the structure at the three dimensional level (von Heijne and Nishikawa, 1991; Bruce, 2001). Plastid transit peptides largely vary in length from an average of 50 up to 146 amino acids (Li and Teng, 2013). There are some features shared with mitochondrial targeting peptides such as the overrepresentation of serine and threonine residues that may explain the targeting of plastid transit peptide containing proteins to mitochondria when expressed in heterologous animal systems (Zhang and Glaser, 2002). No further similarities between plastid and mitochondrial targeting sequences have been identified, and other levels of specificity might exist and enable plant cells to discriminate and accurately sort the two types of organellar proteins. Interestingly, an estimated thirty percent of chloroplast localized proteins do not have a canonical transit peptide (Ferro et al., 2003; Leister, 2003; Kleffmann et al., 2004, 2007; Jarvis, 2008). A recent study in pea indicated that this may be an overestimation that results from a slightly inaccurate algorithm that does not take into account the whole diversity of features of plastid transit peptides (Chang et al., 2014).

The diversity of transit peptides sequences might well be explained by the need to fine tune the import of specific subsets of proteins in agreement with plastid type and developmental stage. Toc159 binds preproteins via their N-terminal, transit peptides (Smith et al., 2004), so one might reasonably expect that the specificity determinants reside within this particular region. However, the determining sequence elements that confer selectivity to a Toc159 isoform have not yet been identified. They could consist of cryptic signals buried in motifs and multiple-motifs (Lee et al., 2009a; Bionda et al., 2010; Chotewutmontri et al., 2012). For example Lee et al. (2009a) revealed that Toc159-dependent import can be mediated by multiple independent motifs, one that consists in a stretch of serine residues located in first 12 amino acid of the N-terminal region of preRBCS (pSSU), and one located in the C-terminal part of the transit peptide sequence (Lee et al., 2009a). In a recent review, (Li and Teng,

2013) analyzed such motifs and their relation with binding sites for various proteins involved in preprotein import. The authors then attributed the preproteins to distinct subgroups based on patterns of sequence motifs in combination with their capacity to be targeted and bind to the protein translocon at the chloroplast outer envelope. Though only a limited number of preproteins were taken into account in these analyses, they clearly indicated that complexity of transit peptide design plays a key role in import selectivity.

REGULATION OF TOC COMPONENTS

EXPRESSION PATTERN

Regulation of TOC complex activity occurs at several levels. Overall the accumulation levels of Toc components throughout development appear to reflect the total import activity, i.e., a highest level of expression for the different components is observed in young, developing tissue, as compared to mature organs (Jarvis et al., 1998; Yu and Li, 2001; Kubis et al., 2003, 2004; Ivanova et al., 2004). As an exception, Toc90 appeared to be uniformly expressed throughout development (Kubis et al., 2003; Infanger et al., 2011). Specific patterns were revealed when comparing the expression levels of the different Toc receptors isoforms in different organs and/or different plastid types, and usually correlated with corresponding mutant phenotypes in *Arabidopsis* (Jarvis et al., 1998; Bauer et al., 2000; Gutensohn et al., 2000; Kubis et al., 2004; Yan et al., 2014). atToc159 and atToc33 are the most highly expressed members of their respective families and both mutants displayed the most severe visible phenotype when compared to other single mutants (Jarvis et al., 1998; Kubis et al., 2004). Furthermore, defects of plastid development in the corresponding mutants follow the expression pattern of the corresponding gene: highly regulated expression is observed for atToc159 and atToc33, with a higher expression occurring in photosynthetic tissues, when compared to other family members. Accordingly single mutants of these genes are specifically affected in plastid type present in those tissues, i.e., the chloroplast and its precursor, the etioplast (Jarvis et al., 1998; Bauer et al., 2000). By the same token, the higher expression of atToc120 and atToc132 in roots correlates with a severe defect of root plastid development in the corresponding double mutant (Kubis et al., 2004). Similarly the mutant phenotype of atToc34, which is expressed more highly in roots, retains normal plastid development but displays reduced root length (Gutensohn et al., 2000; Constan et al., 2004). Thus, selectivity of import into plastids can be modulated at least in part by transcriptional regulation of Toc components in accordance with plant tissue and/or growth conditions (light conditions in the case of Toc159).

Expression profiles of the different Toc members suggest that the receptors acting together in a specific complex are co-regulated at the transcriptional levels. Interestingly, hierarchical cluster analysis indicates that this co-regulation extends to a large variety of conditions (**Figure 2**) and suggests that common *cis* and *trans* regulatory elements could regulate associated Toc receptors. In support of this idea, the CIA2 transcription factor was found to co-modulate atToc33 and atToc75 expression specifically in leaves (Sun et al., 2001, 2009). However, the identity of other transcription factors responsible for the differential expression of Toc members has been poorly investigated so far and further experimentation will be necessary to reveal the molecular mechanisms underlying the regulation of Toc gene expression.

POST-TRANSLATIONAL MODIFICATIONS

Differential regulation of Toc components also occurs at the post-translational levels (**Figure 3**). It is interesting to note that the *ppi2* mutant can be complemented by expression of atToc159 under the constitutive 35S promoter indicating that transcriptional regulation can be bypassed at least under laboratory conditions (Kubis et al., 2004; Agne et al., 2009).

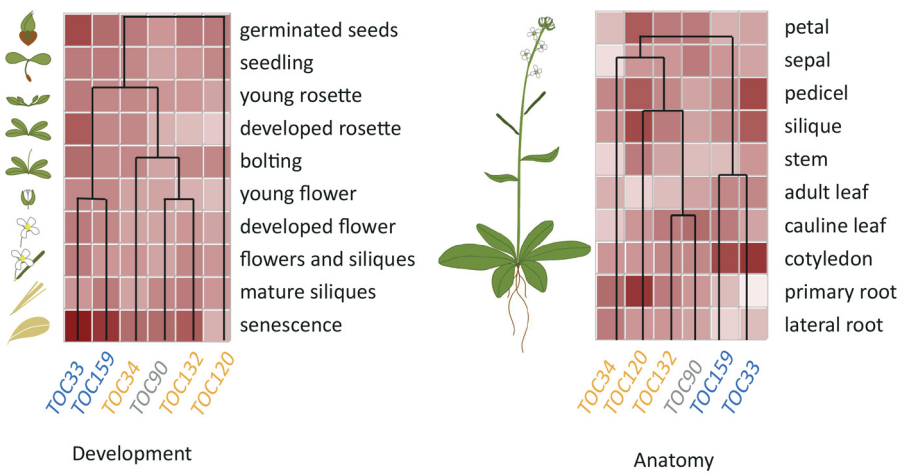


FIGURE 2 | Hierarchical Cluster analysis of *Arabidopsis* Toc receptor gene expression. TOC complexes consist of the assemblies of two different receptors from two separate GTPase families, Toc159/-132/-120/-90 and Toc33/-34, respectively, together with the Toc75 channel. Biochemical and genetic evidence have shown that atToc159 preferentially associates with atToc33 whereas atToc132 and 120 preferentially associate with atToc34. These specific associations are reflected by co-regulation of the Toc receptors isoforms. Data were extracted from Genevestigator database (Nebion), using the Hierarchical Cluster analysis tool, with "Development" or "Anatomy" specific selections for left- and right-hand panels, respectively.

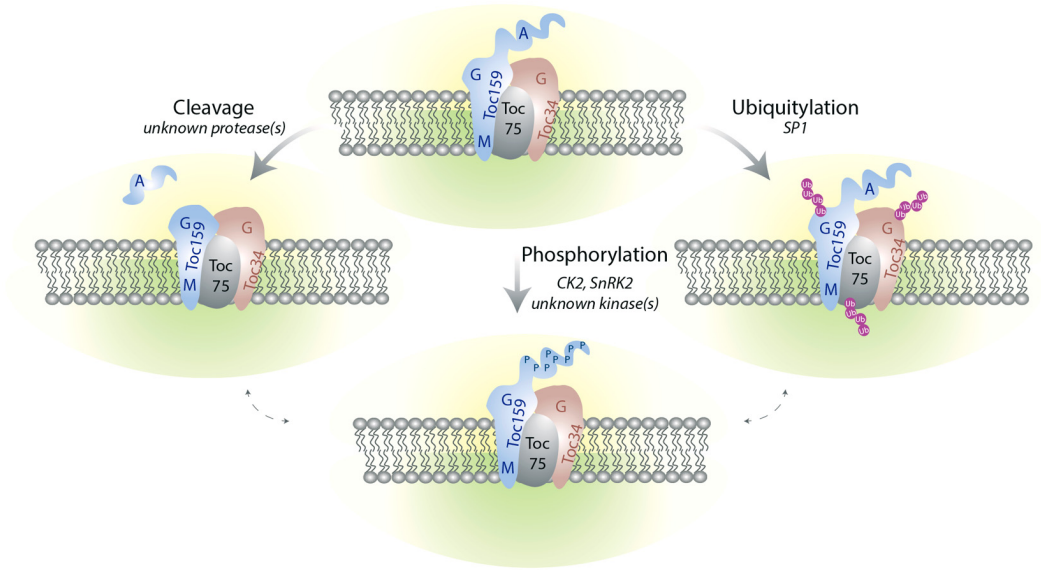


FIGURE 3 | The TOC complex is targeted by multiple post-translational modifications. Phosphorylation of Toc159 and Toc34 at the G-domain may regulate homo and heterodimerization of the Toc receptors as well as their interaction with preproteins. The A-domain of Toc159 is hyperphosphorylated and can be released from the rest of the protein by proteolysis. The functional significance of these A-domain modifications are unclear but they may modulate the selectivity of the receptors for their client preproteins. All the Toc components are subject to ubiquitylation. Ubiquitylation may serve as a signal for proteasome-mediated degradation and pave the way for remodeling of TOC complex composition during plastid differentiation or environmental adaptation. The signaling pathways as well as the environmental and/or developmental factors triggering these PTM remain poorly described. Whether crosstalk between the different types of regulation exists is also not yet known.

Phosphorylation

Several studies have shown that Toc receptors are phosphorylated. Phosphorylation has been reported for pea Toc34 and its ortholog atToc33 (Ser113 and S181, respectively), while it was not detected for atToc34 (Sveshnikova et al., 2000; Fulgosi and Soll, 2002; Jelic et al., 2002, 2003). Differential phosphorylation could therefore represent a regulatory mechanism conferring specificity to the two different members of *Arabidopsis* Toc34 family.

In vitro studies indicated that phosphorylation has a negative effect on GTP and preprotein binding to psToc34 and atToc33 (Sveshnikova et al., 2000; Jelic et al., 2003). Furthermore, *in vitro* and *in vivo* data showed that phosphorylation/phosphomimicking at atToc33 and phosphorylation of psToc34 negatively influenced TOC complex integrity (Oreb et al., 2008). Hypotheses for the underlying molecular mechanisms have been put forward. Since GTPase activity may be required for G-domain-mediated association of Toc159 and Toc34 (Smith et al., 2002; Wallas et al., 2003), phosphorylation may indirectly prevent homo- as well as heterodimerization because of a negative effect on GTP-binding. More directly the bulky, negatively charged phosphate group could inhibit the binding to a preprotein or to Toc159. However, this latter hypothesis may be valid for *Arabidopsis*, but not for pea since the phosphorylation site is distant from the dimerization interface (Oreb et al., 2008). In summary, the available data suggest the phosphorylation of psToc34 and atToc33 have a dual function, regulating both TOC complex assembly and subsequent substrate binding.

The physiological relevance and the signals triggering this specific phosphorylation are still not clearly defined. Data obtained from *Arabidopsis* transgenic lines expressing phosphomicking variants of atToc33 confirmed that phosphorylation at S181 can inhibit atToc33 activity in young *Arabidopsis* seedlings but not later during development (Aronsson et al., 2006; Oreb et al., 2007). Indeed, phosphomimick variants resemble the *ppi1* mutant regarding a number of phenotypic traits in 5 day-old *Arabidopsis* seedlings (chlorophyll accumulation, chloroplast ultrastructure, and photosynthetic activity). However, since the non-phosphorylatable version behaved similarly to the WT, it was not possible to determine the conditions under which atToc33 is phosphorylated in *planta* (Aronsson et al., 2006; Oreb et al., 2007). We speculate that phosphorylation might represent a means to quickly down-regulate preprotein import *via* atToc33 containing TOC complexes, for example in mature plastids where protein demand is low. Moreover and since atToc33 can be phosphorylated but not atToc34, this post-translational regulation may affect the selectivity aspect of preprotein import regulation.

One additional phosphorylation site has been experimentally identified in both atToc33 and -34 [data provided by PhosphAT (Durek et al., 2010)]. It maps to a conserved Tyrosine residue of the G-domain. Additional studies will be required to validate and determine the regulatory effect of this specific phosphorylation.

Finally, the identity of Toc33/Toc34 kinase(s) still remain(s) mysterious. Some clues stemming from pea suggest that psToc34 is phosphorylated by an ATP-dependent, 98 kDa kinase residing at the outer envelope membrane (Fulgosi and Soll, 2002).

However, the amino acid sequence information is not sufficient to molecularly identify the potential kinase in pea or its homolog in *Arabidopsis*.

The Toc159 receptors are also targets of phosphorylation. First evidence of phosphorylation of Toc159 came from *in vitro* studies using outer envelopes isolated from pea chloroplasts, showing that both full length Toc159 and its natural 86 kDa fragment could be phosphorylated (Fulgosi and Soll, 2002). Phosphorylation was demonstrated for the G-domain of psToc159, reminiscent of Toc33/34 regulation (Oreb et al., 2008), however, neither the precise site nor the regulatory function were further investigated. Large-scale phosphoproteomics projects revealed that Toc159 members in *Arabidopsis* are highly phosphorylated at the acidic A-domain (Agne et al., 2010; Durek et al., 2010). In total, 43 sites have been mapped in atToc159, while far fewer were detected in the other three members. These lower numbers may be due to the shorter length of the atToc132 and atToc120 A-domains, the absence of such a domain in atToc90, or because lower protein accumulation levels when compared to atToc159 limit the detection by mass spectrometry. Nevertheless the identified phosphorylation sites do not map to matching positions in the different homologs, which confers an additional degree of divergence to the A-domain.

The functional relevance of A-domain phosphorylation has been poorly documented so far. The dispensable nature of the A-domain suggests that phosphorylation either plays a minor role altogether, or possibly an important regulatory role under specific conditions (Hiltbrunner et al., 2001b; Agne et al., 2009; Inoue et al., 2010). The A-domain behaves as an intrinsically disordered protein, which is often linked to multiple and transient protein–protein interactions (Richardson et al., 2009). Therefore phosphorylation of this domain could modulate interactions of Toc159 with other Toc components but also with specific sets of client preproteins. In addition, a selective autoinhibitory function of the A-domain under specific conditions may be envisaged that may be alleviated by phosphorylation or proteolytic removal.

Recently a link between ABA signaling and phosphorylation of Toc159 family members in *Arabidopsis* has been established (Wang et al., 2013). Upon ABA treatment atToc159 was phosphorylated at Thr692. atToc120 and atToc132 phosphopeptides accumulation was also enhanced by ABA treatment. These data together with the fact that a mutant deficient in ABA synthesis is affected in pre-protein import and early plant development suggest a close link between ABA signaling and chloroplast protein import regulation via Toc159 A-domain phosphorylation (Zhong et al., 2010). Whether ABA dependent phosphorylation plays a role in preprotein recognition, impacts TOC159 complex assembly, or acts at the level of the translocation process will be interesting questions to be addressed in the future.

Several classes of kinases may mediate phosphorylation of Toc159 homologs. Motif analysis suggests that a large fraction of atToc159 phosphorylation sites represent potential cytosolic casein kinase 2 (CK2) targets and this was validated biochemically by *in vitro* phosphorylation experiments (Agne et al., 2010). Recently it has been shown that ABA dependent phosphorylation of atToc159 at Thr692 was decreased in a triple mutant *snrk2.2/2.3/2.6* that is nearly insensitive to ABA treatment (Wang et al., 2013). In addition SnRK2.6 phosphorylated recombinant atToc159 *in vitro*. Thus SnRK2.6 represents a potential kinase of atToc159 at Thr692. On the contrary, atToc120 and atToc132 phosphorylation upon ABA treatment was detected only in the triple mutant *snrk2.2/2.3/2.6*, indicating the involvement of another ABA regulated kinase. Indeed ABA signaling is mediated by multiple kinases of the SnRK family but also of the MAPK kinase family (Danquah et al., 2014). The phosphorylation status of Toc159 members could therefore be regulated antagonistically by ABA signaling via the action of different classes of kinases and could represent a way to switch between Toc132/Toc120 and Toc159 specific import depending on environmental as well as developmental conditions and consequent plastid preprotein requirements. Finally, it has been proposed that psToc159 is a target of a 70 kDa kinase located at the outer envelope of the pea chloroplast (Fulgosi and Soll, 2002) but so far no study has reported on the identification of a putative homolog in *Arabidopsis*.

In conclusion phosphorylation of the Toc159 and Toc34 receptors potentially regulates protein import at different levels: it may impact the import rate by regulating the affinity toward client preproteins, or affect the composition of the TOC complex by modulating the interaction between Toc receptors and consequently change the selectivity of plastid protein import. The involvement of ABA signaling in this regulation indicates that phosphorylation of Toc components can modulate the import activity in response to developmental signals for example during germination or subsequent post-germinative processes, or in response to abiotic stress that require the tuning of the plastid proteome. Hormonal control of plastid development has been frequently reported, but the effects on import activity are still poorly documented.

Phosphorylation could also be part of a signaling cascade enabling subsequent additional post-translational modifications (PTM) since cross talk between different is a common phenomenon in eukaryotic systems, and PTM other than phosphorylation have been described for the different Toc components (see below). The existence of numerous phosphorylation sites, especially in Toc159 families, suggests the participation of multiple kinases, and corresponding signaling pathways probably acting in a network.

Post-translational modifications other than phosphorylation

Toc159 was first identified as an 86 kDa protein lacking the A-domain (Hirsch et al., 1994; Kessler et al., 1994; Schnell et al., 1994; Bolter et al., 1998). It is not clear whether proteolysis occurs only during chloroplast preparation or whether it is part of regulatory system acting on Toc159. It is not clear either if other Toc159 homologs are also substrates of proteolytic cleavage but the relative stability of the A-domain fragment of atToc159 favors controlled proteolysis (Agne et al., 2010). Therefore, a yet unknown protease may process Toc159 conditionally, leading to the removal of the A-domain and consequently altering the import selectivity. Interplay between phosphorylation and cleavage has been demonstrated in other biological systems for example in the context of apoptosis (Dix et al., 2012). Investigation of the cross talk between these two PTM will certainly be an interesting aspect for future research.

Abundance of the different Toc members varies developmentally. Currently an important question is to understand how the TOC machinery is remodeled upon plastid development and plastid inter-conversion. As discussed above transcriptional regulation plays a role in modulation of Toc components expression depending on plant tissues and environmental conditions, while PTM may participate in the regulation of TOC complex assembly and activity. Recently a genetic study complemented by biochemical analyses revealed that Toc receptors as well as the Toc75 channel could be modified by ubiquitylation. Ubiquitylation required SP1, a chloroplast outer membrane localized E3 ubiquitin ligase (Ling et al., 2012). Enhanced accumulation of TOC proteins in *sp1* genetic background suggested that SP1 indeed participates in UPS-mediated degradation of Toc components. Phenotypic analyses indicated that this regulatory mechanism may play a role during plastid inter-conversion. However, how SP1 is regulated and functions selectively on the different Toc receptors has not been addressed so far. Again a possible interplay with phosphorylation regulation might be envisaged as phosphorylation can serve as either a positive or a negative signal for ubiquitylation (Hunter, 2007).

CONCLUDING REMARKS

Acquisition of the capacity to target proteins to different compartments has enabled eukaryotic cells to maintain and control the development of organelles. In higher plants the evolution of the TOC–TIC machinery has been a key mechanism enabling developmental processes. The evolutionary diversification of Toc receptors and transit peptides likely led to the tissue- and plastid type dependent preprotein selectivity of the import process. It is now well accepted that preprotein import in plastids plays a central role in the maintenance of cellular homeostasis, controlling the development and differentiation of this organelle. In a more indirect way, preprotein import also exerts control of nuclear gene expression via retrograde signaling to the nucleus. The composition and mode of action of the import machinery has been studied extensively in the past years, and now progress needs to be made toward the understanding of the regulatory mechanisms controlling the assembly and the activity of the complex. Regulation is not only important for correct sorting of preproteins, but also to limit energy expenditure associated with this costly process. Multiple types of PTM of Toc receptors have been discovered; however, their functional significance largely remains in the dark. Identification of the regulatory factors and signaling pathways as well as unraveling the biological relevance of the various PTM at the import machinery will provide new insight on how plants control development and adapt to the environment.

ACKNOWLEDGMENTS

This work was supported by UniNE, SNF31003A_127380 and SNF31003A_144156, and Marie Heim Voegtlin PMPDP3_151301.

REFERENCES

Agne, B., Andres, C., Montandon, C., Christ, B., Ertan, A., Jung, F., et al. (2010). The acidic A-domain of *Arabidopsis* TOC159 occurs as a hyperphosphorylated protein. *Plant Physiol.* 153, 1016–1030. doi: 10.1104/pp.110.158048

Agne, B., Infanger, S., Wang, F., Hofstetter, V., Rahim, G., Martin, M., et al. (2009). A toc159 import receptor mutant, defective in hydrolysis of GTP, supports preprotein import into chloroplasts. *J. Biol. Chem.* 284, 8670–8679. doi: 10.1074/jbc.M804235200

Andres, C., Agne, B., and Kessler, F. (2010). The TOC complex: preprotein gateway to the chloroplast. *Biochim. Biophys. Acta* 1803, 715–723. doi: 10.1016/j.bbamcr.2010.03.004

Aronsson, H., Combe, J., Patel, R., Agne, B., Martin, M., Kessler, F., et al. (2010). Nucleotide binding and dimerization at the chloroplast pre-protein import receptor, atToc33, are not essential in vivo but do increase import efficiency. *Plant J.* 63, 297–311. doi: 10.1111/j.1365-313X.2010.04242.x

Aronsson, H., Combe, J., Patel, R., and Jarvis, P. (2006). In vivo assessment of the significance of phosphorylation of the *Arabidopsis* chloroplast protein import receptor, atToc33. *FEBS lett.* 580, 649–655. doi: 10.1016/j.febslet.2005.12.055

Aronsson, H., and Jarvis, P. (2011). Dimerization of TOC receptor GTPases and its implementation for the control of protein import into chloroplasts. *Biochem. J.* 436, e1–e2. doi: 10.1042/BJ20110659

Asano, T., Yoshioka, Y., and Machida, Y. (2004). A defect in atToc159 of *Arabidopsis thaliana* causes severe defects in leaf development. *Genes Genet. Syst.* 79, 207–212. doi: 10.1266/ggs.79.207

Barsan, C., Zouine, M., Maza, E., Bian, W., Egea, I., Rossignol, 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. doi: 10.1104/pp.112.203679

Bauer, J., Chen, K., Hiltbunner, A., Wehrli, E., Eugster, M., Schnell, D., et al. (2000). The major protein import receptor of plastids is essential for chloroplast biogenesis. *Nature* 403, 203–207. doi: 10.1038/35003214

Becker, T., Jelic, M., Vojta, A., Radunz, A., Soll, J., and Schleiff, E. (2004). Preprotein recognition by the Toc complex. *EMBO J.* 23, 520–530. doi: 10.1038/sj.emboj.7600089

Bionda, T., Tillmann, B., Simm, S., Beilstein, K., Ruprecht, M., and Schleiff, E. (2010). Chloroplast import signals: the length requirement for translocation in vitro and in vivo. *J. Mol. Biol.* 402, 510–523. doi: 10.1016/j.jmb.2010.07.052

Bischof, S., Baerenfaller, K., Wildhaber, T., Troesch, R., Vidi, P. A., Roschitzki, B., et al. (2011). Plastid proteome assembly without Toc159: photosynthetic protein import and accumulation of N-acetylated plastid precursor proteins. *Plant Cell* 23, 3911–3928. doi: 10.1105/tpc.111.092882

Blobel, G. (1980). Regulation of intracellular protein traffic. *Harvey Lect.* 76, 125–147.

Bolter, B., May, T., and Soll, J. (1998). A protein import receptor in pea chloroplasts, Toc86, is only a proteolytic fragment of a larger polypeptide. *FEBS Lett.* 441, 59–62. doi: 10.1016/S0014-5793(98)01525-7

Brautigam, A., and Weber, A. P. (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. doi: 10.1093/mp/ssp070

Bruce, B. D. (2001). The paradox of plastid transit peptides: conservation of function despite divergence in primary structure. *Biochim. Biophys. Acta* 1541, 2–21. doi: 10.1016/S0167-4889(01)00149-5

Chang, W., Soll, J., and Bolter, B. (2014). A new member of the psToc159 family contributes to distinct protein targeting pathways in pea chloroplasts. *Front. Plant Sci.* 5:239. doi: 10.3389/fpls.2014.00239

Chotewutmontri, P., Reddick, L. E., McWilliams, D. R., Campbell, I. M., and Bruce, B. D. (2012). Differential transit peptide recognition during preprotein binding and translocation into flowering plant plastids. *Plant Cell* 24, 3040–3059. doi: 10.1105/tpc.112.098327

Constan, D., Patel, R., Keegstra, K., and Jarvis, P. (2004). An outer envelope membrane component of the plastid protein import apparatus plays an essential role in *Arabidopsis*. *Plant J.* 38, 93–106. doi: 10.1111/j.1365-313X.2004.02024.x

Dahlin, C., and Cline, K. (1991). Developmental regulation of the plastid protein import apparatus. *Plant Cell* 3, 1131–1140. doi: 10.1105/tpc.3.10.1131

Danquah, A., de Zelicourt, A., Colcombet, J., and Hirt, H. (2014). The role of ABA and MAPK signaling pathways in plant abiotic stress responses. *Biotechnol. Adv.* 32, 40–52. doi: 10.1016/j.biotechadv.2013.09.006

Dix, M. M., Simon, G. M., Wang, C., Okerberg, E., Patricelli, M. P., and Cravatt, B. F. (2012). Functional interplay between caspase cleavage and phosphorylation sculpts the apoptotic proteome. *Cell* 150, 426–440. doi: 10.1016/j.cell.2012.05.040

Durek, P., Schmidt, R., Heazlewood, J. L., Jones, A., MacLean, D., Nagel, A., et al. (2010). PhosPhAt: the *Arabidopsis thaliana* phosphorylation site database. An update. *Nucleic Acids Res.* 38, D828–D834. doi: 10.1093/nar/gkp810

Dutta, S., Mohanty, S., and Tripathy, B. C. (2009). Role of temperature stress on chloroplast biogenesis and protein import in pea. *Plant Physiol.* 150, 1050–1061. doi: 10.1104/pp.109.137265

Dutta, S., Teresinski, H. J., and Smith, M. D. (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. doi: 10.1371/journal.pone.0095026

Ferro, M., Salvi, D., Brugiere, S., Miras, S., Kowalski, S., Louwagie, M., et al. (2003). Proteomics of the chloroplast envelope membranes from *Arabidopsis thaliana*. *Mol. Cell Proteomics* 2, 325–345.

Friso, G., Giacomelli, L., Ytterberg, A. J., Peltier, J. B., Rudella, A., Sun, Q., et al. (2004). In-depth analysis of the thylakoid membrane proteome of *Arabidopsis thaliana* chloroplasts: new proteins, new functions, and a plastid proteome database. *Plant cell* 16, 478–499. doi: 10.1105/tpc.017814

Fulgosi, H., and Soll, J. (2002). The chloroplast protein import receptors Toc34 and Toc159 are phosphorylated by distinct protein kinases. *J. Biol. Chem.* 277, 8934–8940. doi: 10.1074/jbc.M110679200

Gagat, P., Bodyl, A., and Mackiewicz, P. (2013). How protein targeting to primary plastids via the endomembrane system could have evolved? A new hypothesis based on phylogenetic studies. *Biol. direct* 8, 18. doi: 10.1186/1745-6150-8-18

Gross, J., and Bhattacharya, D. (2009). Revaluating the evolution of the Toc and Tic protein translocons. *Trends Plant Sci.* 14, 13–20. doi: 10.1016/j.tplants.2008.10.003

Gutensohn, M., Schulz, B., Nicolay, P., and Flugge, U. I. (2000). Functional analysis of the two *Arabidopsis* homologues of Toc34, a component of the chloroplast protein import apparatus. *Plant J.* 23, 771–783. doi: 10.1046/j.1365-313x.2000.00849.x

Hiltbrunner, A., Bauer, J., Alvarez-Huerta, M., and Kessler, F. (2001a). Protein translocon at the *Arabidopsis* outer chloroplast membrane. *Biochem. Cell Biol.* 79, 629–635. doi: 10.1139/bcb-79-5-629

Hiltbrunner, A., Bauer, J., Vidi, P. A., Infanger, S., Weibel, P., Hohwy, M., et al. (2001b). Targeting of an abundant cytosolic form of the protein import receptor at Toc159 to the outer chloroplast membrane. *J. Cell Biol.* 154, 309–316. doi: 10.1083/jcb.200104022

Hirsch, S., Muckel, E., Heemeyer, F., von Heijne, G., and Soll, J. (1994). A receptor component of the chloroplast protein translocation machinery. *Science* 266, 1989–1992. doi: 10.1126/science.7801125

Hunter, T. (2007). The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Mol. Cell* 28, 730–738. doi: 10.1016/j.molcel.2007.11.019

Inaba, T., Alvarez-Huerta, M., Li, M., Bauer, J., Ewers, C., Kessler, F., et al. (2005). *Arabidopsis* tic110 is essential for the assembly and function of the protein import machinery of plastids. *Plant Cell* 17, 1482–1496. doi: 10.1105/tpc.105.030700

Infanger, S., Bischof, S., Hiltbrunner, A., Agne, B., Baginsky, S., and Kessler, F. (2011). The chloroplast import receptor Toc90 partially restores the accumulation of Toc159 client proteins in the *Arabidopsis thaliana* ppi2 mutant. *Mol. plant* 4, 252–263. doi: 10.1093/mp/ssq071

Inoue, H., Rounds, C., and Schnell, D. J. (2010). The molecular basis for distinct pathways for protein import into *Arabidopsis* chloroplasts. *Plant Cell* 22, 1947–1960. doi: 10.1105/tpc.110.074328

Inoue, K., and 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. doi: 10.1046/j.1365-313X.2003.01755.x

Ivanova, Y., Smith, M. D., Chen, K., and Schnell, D. J. (2004). Members of the Toc159 import receptor family represent distinct pathways for protein targeting to plastids. *Mol. Biol. Cell* 15, 3379–3392. doi: 10.1091/mbc.E03-12-0923

Jackson-Constan, D., and Keegstra, K. (2001). *Arabidopsis* genes encoding components of the chloroplastic protein import apparatus. *Plant Physiol.* 125, 1567–1576. doi: 10.1104/pp.125.4.1567

Jarvis, P. (2008). Targeting of nucleus-encoded proteins to chloroplasts in plants. *New Phytol.* 179, 257–285. doi: 10.1111/j.1469-8137.2008.02452.x

Jarvis, P., Chen, L. J., Li, H., Peto, C. A., Fankhauser, C., and Chory, J. (1998). An *Arabidopsis* mutant defective in the plastid general protein import apparatus. *Science* 282, 100–103. doi: 10.1126/science.282.5386.100

Jarvis, P., and Lopez-Juez, E. (2013). Biogenesis and homeostasis of chloroplasts and other plastids. *Nat. Rev. Mol. Cell Biol.* 14, 787–802. doi: 10.1038/nrm3702

Jelic, M., Soll, J., and Schleiff, E. (2003). Two Toc34 homologues with different properties. *Biochemistry* 42, 5906–5916. doi: 10.1021/bi034001q

Jelic, M., Sveshnikova, N., Motzkus, M., Horth, P., Soll, J., and Schleiff, E. (2002). The chloroplast import receptor Toc34 functions as preprotein-regulated GTPase. *Biol. Chem.* 383, 1875–1883. doi: 10.1515/BC.2002.211

Jensen, P. E., and Leister, D. (2014). Chloroplast evolution, structure and functions. *F1000Prime Rep.* 6, 40. doi: 10.12703/P6-40

Kakizaki, T., Matsumura, H., Nakayama, K., Che, F. S., Terauchi, R., and Inaba, T. (2009). Coordination of plastid protein import and nuclear gene expression by plastid-to-nucleus retrograde signaling. *Plant Physiol.* 151, 1339–1353. doi: 10.1104/pp.109.145987

Kalanon, M., and McFadden, G. I. (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. doi: 10.1534/genetics.107.085704

Kessler, F., Blobel, G., Patel, H. A., and Schnell, D. J. (1994). Identification of two GTP-binding proteins in the chloroplast protein import machinery. *Science* 266, 1035–1039. doi: 10.1126/science.7973656

Kessler, F., and Schnell, D. J. (2002). A GTPase gate for protein import into chloroplasts. *Nat. Struct. Biol.* 9, 81–83. doi: 10.1038/nsb0202-81

Kessler, F., and Schnell, D. J. (2006). The function and diversity of plastid protein import pathways: a multilane GTPase highway into plastids. *Traffic* 7, 248–257. doi: 10.1111/j.1600-0854.2005.00382.x

Kikuchi, S., Bedard, J., Hirano, M., Hirabayashi, Y., Oishi, M., Imai, M., et al. (2013). Uncovering the protein translocon at the chloroplast inner envelope membrane. *Science* 339, 571–574. doi: 10.1126/science.1229262

Kikuchi, S., Oishi, M., Hirabayashi, Y., Lee, D. W., Hwang, I., and Nakai, M. (2009). A 1-megadalton translocation complex containing Tic20 and Tic21 mediates chloroplast protein import at the inner envelope membrane. *Plant Cell* 21, 1781–1797. doi: 10.1105/tpc.108.063552

Kleffmann, T., Russenberger, D., von Zychlinski, A., Christopher, W., Sjolander, K., Gruissem, W., et al. (2004). The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. *Curr. Biol.* 14, 354–362. doi: 10.1016/j.cub.2004.02.039

Kleffmann, T., von Zychlinski, A., Russenberger, D., Hirsch-Hoffmann, M., Gehrig, P., Gruissem, W., et al. (2007). Proteome dynamics during plastid differentiation in rice. *Plant Physiol.* 143, 912–923. doi: 10.1104/pp.106.090738

Kovacheva, S., Bedard, J., Patel, R., Dudley, P., Twell, D., Rios, G., et al. (2005). In vivo studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. *Plant J.* 41, 412–428. doi: 10.1111/j.1365-313X.2004.02307.x

Kovacs-Bogdan, E., Benz, J. P., Soll, J., and Bolter, B. (2011). Tic20 forms a channel independent of Tic110 in chloroplasts. *BMC Plant Biol.* 11:133. doi: 10.1186/1471-2229-11-133

Kovacs-Bogdan, E., Soll, J., and Bolter, B. (2010). Protein import into chloroplasts: the Tic complex and its regulation. *Biochim. Biophys. Acta* 1803, 740–747. doi: 10.1016/j.bbamcr.2010.01.015

Kubis, S., Baldwin, A., Patel, R., Razzaq, A., Dupree, P., Lilley, K., et al. (2003). The *Arabidopsis* ppi1 mutant is specifically defective in the expression, chloroplast import, and accumulation of photosynthetic proteins. *Plant Cell* 15, 1859–1871. doi: 10.1105/tpc.012955

Kubis, S., Patel, R., Combe, J., Bedard, J., Kovacheva, S., Lilley, K., et al. (2004). Functional specialization amongst the *Arabidopsis* Toc159 family of chloroplast protein import receptors. *Plant Cell* 16, 2059–2077. doi: 10.1105/tpc.104.023309

Lee, D. W., Lee, S., Oh, Y. J., and Hwang, I. (2009a). Multiple sequence motifs in the rubisco small subunit transit peptide independently contribute to Toc159-dependent import of proteins into chloroplasts. *Plant Physiol.* 151, 129–141. doi: 10.1104/pp.109.140673

Lee, J., Wang, F., and Schnell, D. J. (2009b). Toc receptor dimerization participates in the initiation of membrane translocation during protein import into chloroplasts. *J. Biol. Chem.* 284, 31130–31141. doi: 10.1074/jbc.M109.053751

Leister, D. (2003). Chloroplast research in the genomic age. *Trends Genet.* 19, 47–56. doi: 10.1016/S0168-9525(02)00003-3

Li, H. M., and Chiu, C. C. (2010). Protein transport into chloroplasts. *Annu. Rev. Plant Biol.* 61, 157–180. doi: 10.1146/annurev-arplant-042809-112222

Li, H. M., and Teng, Y. S. (2013). Transit peptide design and plastid import regulation. *Trends Plant Sci.* 18, 360–366. doi: 10.1016/j.tplants.2013.04.003

Ling, Q., Huang, W., Baldwin, A., and Jarvis, P. (2012). Chloroplast biogenesis is regulated by direct action of the ubiquitin-proteasome system. *Science* 338, 655–659. doi: 10.1126/science.1225053

Olsen, L. J., and Keegstra, K. (1992). The binding of precursor proteins to chloroplasts requires nucleoside triphosphates in the intermembrane space. *J. Biol. Chem.* 267, 433–439.

Oreb, M., Hofle, A., Koenig, P., Sommer, M. S., Sinning, I., Wang, F., et al. (2011). Substrate binding disrupts dimerization and induces nucleotide exchange of the chloroplast GTPase Toc33. *Biochem. J.* 436, 313–319. doi: 10.1042/BJ20110246

Oreb, M., Hofle, A., Mirus, O., and Schleiff, E. (2008). Phosphorylation regulates the assembly of chloroplast import machinery. *J. Exp. Bot.* 59, 2309–2316. doi: 10.1093/jxb/ern095

Oreb, M., Zoryan, M., Vojta, A., Maier, U. G., Eichacker, L. A., and Schleiff, E. (2007). Phospho-mimicry mutant of atToc33 affects early development of *Arabidopsis thaliana*. *FEBS Lett.* 581, 5945–5951. doi: 10.1016/j.febslet.2007.11.071

Perry, S. E., and Keegstra, K. (1994). Envelope membrane proteins that interact with chloroplastic precursor proteins. *Plant Cell* 6, 93–105. doi: 10.1105/tpc.6.1.93

Primavesi, L. F., Wu, H., Mudd, E. A., Day, A., and Jones, H. D. (2008). Visualisation of plastids in endosperm, pollen and roots of transgenic wheat expressing modified GFP fused to transit peptides from wheat SSU RubisCO, rice FtsZ and maize ferredoxin III proteins. *Transgenic Res.* 17, 529–543. doi: 10.1007/s11248-007-9126-7

Rahim, G., Bischof, S., Kessler, F., and Agne, B. (2009). In vivo interaction between atToc33 and atToc159 GTP-binding domains demonstrated in a plant split-ubiquitin system. *J. Exp. Bot.* 60, 257–267. doi: 10.1093/jxb/ern283

Reumann, S., Inoue, K., and Keegstra, K. (2005). Evolution of the general protein import pathway of plastids (review). *Mol. Membr. Biol.* 22, 73–86. doi: 10.1080/09687860500041916

Richardson, L. G., Jelokhani-Niaraki, M., and Smith, M. D. (2009). The acidic domains of the Toc159 chloroplast preprotein receptor family are intrinsically disordered protein domains. *BMC Biochem.* 10:35. doi: 10.1186/1471-2091-10-35

Schleiff, E., and Soll, J. (2005). Membrane protein insertion: mixing eukaryotic and prokaryotic concepts. *EMBO Rep* 6, 1023–1027. doi: 10.1038/sj.embor.7400563

Schnell, D. J., Kessler, F., and Blobel, G. (1994). Isolation of components of the chloroplast protein import machinery. *Science* 266, 1007–1012. doi: 10.1126/science.7973649

Shi, L. X., and Theg, S. M. (2013). The chloroplast protein import system: from algae to trees. *Biochim. Biophys. Acta* 1833, 314–331. doi: 10.1016/j.bbamcr.2012.10.002

Smith, D. R., and Lee, R. W. (2014). A plastid without a genome: evidence from the nonphotosynthetic green algal genus Polytomella. *Plant Physiol.* 164, 1812–1819. doi: 10.1104/pp.113.233718

Smith, M. D., Hiltbrunner, A., Kessler, F., and Schnell, D. J. (2002). The targeting of the atToc159 preprotein receptor to the chloroplast outer membrane is mediated by its GTPase domain and is regulated by GTP. *J. Cell Biol.* 159, 833–843. doi: 10.1083/jcb.200208017

Smith, M. D., Rounds, C. M., Wang, F., Chen, K., Afitlhile, M., and Schnell, D. J. (2004). atToc159 is a selective transit peptide receptor for the import of nucleus-encoded chloroplast proteins. *J. Cell Biol.* 165, 323–334. doi: 10.1083/jcb.200311074

Sun, C. W., Chen, L. J., Lin, L. C., and Li, H. M. (2001). Leaf-specific upregulation of chloroplast translocon genes by a CCT motif-containing protein, CIA 2. *Plant Cell* 13, 2053–2061.

Sun, C. W., Huang, Y. C., and Chang, H. Y. (2009). CIA2 coordinately up-regulates protein import and synthesis in leaf chloroplasts. *Plant Physiol.* 150, 879–888. doi: 10.1104/pp.109.137240

Sun, Y. J., Forouhar, F., Li Hm, H. M., Tu, S. L., Yeh, Y. H., Kao, S., et al. (2002). Crystal structure of pea Toc34, a novel GTPase of the chloroplast protein translocon. *Nat. Struct. Biol.* 9, 95–100. doi: 10.1038/nsb744

Sveshnikova, N., Soll, J., and Schleiff, E. (2000). Toc34 is a preprotein receptor regulated by GTP and phosphorylation. *Proc. Natl. Acad. Sci. U.S.A.* 97, 4973–4978. doi: 10.1073/pnas.080491597

Teng, Y. S., Chan, P. T., and Li, H. M. (2012). Differential age-dependent import regulation by signal peptides. *PLoS Biol.* 10:e1001416. doi: 10.1371/journal.pbio.1001416

von Heijne, G., and Nishikawa, K. (1991). Chloroplast transit peptides. the perfect random coil? *FEBS lett.* 278, 1–3. doi: 10.1016/0014-5793(91)80069-F

Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M., and Tommassen, J. (2003). Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* 299, 262–265. doi: 10.1126/science.1078973

Wallas, T. R., Smith, M. D., Sanchez-Nieto, S., and Schnell, D. J. (2003). The roles of toc34 and toc75 in targeting the toc159 preprotein receptor to chloroplasts. *J. Biol. Chem.* 278, 44289–44297. doi: 10.1074/jbc.M307873200

Wan, J., Blakeley, S. D., Dennis, D. T., and 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. doi: 10.1074/jbc.271.49.31227

Wang, P., Xue, L., Batelli, G., Lee, S., Hou, Y. J., Van Oosten, M. J., et al. (2013). Quantitative phosphoproteomics identifies SnRK2 protein kinase substrates and reveals the effectors of abscisic acid action. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11205–11210. doi: 10.1073/pnas.1308974110

Yan, J., Campbell, J. H., Glick, B. R., Smith, M. D., and Liang, Y. (2014). Molecular characterization and expression analysis of chloroplast protein import components in tomato (*Solanum lycopersicum*). *PLoS ONE* 9:e95088. doi: 10.1371/journal.pone.0095088

Yu, T. S., and 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. doi: 10.1104/pp.127.1.90

Zhang, X. P., and Glaser, E. (2002). Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone. *Trends Plant Sci.* 7, 14–21. doi: 10.1016/S1360-1385(01)02180-X

Zhong, R., Thompson, J., Ottesen, E., and Lamppa, G. K. (2010). A forward genetic screen to explore chloroplast protein import in vivo identifies Moco sulfurase, pivotal for ABA and IAA biosynthesis and purine turnover. *Plant J.* 63, 44–59. doi: 10.1111/j.1365-313X.2010.04220.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 17 July 2014; accepted: 01 September 2014; published online: 17 September 2014.

*Citation: Demarsy E, Lakshmanan AM and Kessler F (2014) Border control: selectivity of chloroplast protein import and regulation at the TOC-complex. Front. Plant Sci. **5**:483. doi: 10.3389/fpls.2014.00483*

This article was submitted to Plant Cell Biology, a section of the journal Frontiers in Plant Science.