8-2016

Decoding the cellular *zipcode*: Functional analysis of transit peptide motifs and mechanistic implications in plastid targeting and import

Kristen N. Holbrook
University of Tennessee, Knoxville, kholbrook@tennessee.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Kristen N. Holbrook entitled "Decoding the cellular *zipcode*: Functional analysis of transit peptide motifs and mechanistic implications in plastid targeting and import." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Biochemistry and Cellular and Molecular Biology.

Barry D. Bruce, Major Professor

We have read this dissertation and recommend its acceptance:

Andreas Nebenfuehr, Francisco Barrera, Xiaolin Cheng, Todd Reynolds

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Decoding the cellular *zipcode*: Functional analysis of transit peptide motifs and mechanistic implications in plastid targeting and import

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Kristen N. Holbrook
August 2016

ACKNOWLEDGEMENTS

I am grateful for all of the people who have helped me with my research project over the years, most of all to my advisor, Dr. Barry Bruce, who accepted me into his lab, believed in me, and provided me with the opportunity to develop a career in research science. I would like to thank my doctoral committee, Drs. Andreas Nebenfuehr, Francisco Barrera, Xiaolin Cheng and Todd Reynolds for their help and support. Drs. Albrecht von Arnim and Andreas Nebenfuehr provided biolistic equipment and the microscope for my experiments. Dr. John Dunlap assisted me with confocal microscopy. I would like to thank the organizations that supported my travel and research in graduate school, including the National Science Foundation, Gibson Family Foundation, Tennessee Plant Research Center, and the BCMB department. Finally, I would like to thank my amazing friends and labmates for their support, encouragement, and many laughs over the years: especially Khoa Nguyen, Meng Li, Nate Brady and Jyotirmoy Mondal. This project was not possible without the contributions of talented undergraduates over the years, many thanks to Amber Bassett, Erika Sanders, and Rena Abdurehman for their time and efforts.

ABSTRACT

Eukaryotic organisms are defined by their compartmentalization and various organelles. The membranes that define these organelles require complex nanomachines (known as translocons) to selectively mediate the import of proteins from the cytosol where they are synthesized into the organelle. The plastid, (specifically the chloroplast) which is characteristic of plant cells, possibly represents the most complex system of protein sorting, requiring many different translocons located in the three membranes found in this organelle. Despite having a small genome, the vast majority of plastid-localized proteins are nuclear-encoded and must be post-translationally imported from the cytosol. These proteins are encoded as a larger molecular weight precursor that contains a special “zip code” at their N-terminus appropriately called a *transit peptide* (TP).

The mechanism of how the transit peptide mediates the selective targeting and translocation of precursor proteins into chloroplasts remains unclear. We aim to precisely understand how transit peptides are recognized and processed by the Translocon of the Outer Chloroplast membrane (abbreviated TOC). This is still a paradox, since several thousand different transit peptides exist in a given plant species, yet show very little similarity with each other – however, they seem to function in a common pathway that involves the recognition one or more receptor GTPases, which are proposed to function as “gatekeepers.” The current study extends our understanding of both TP physicochemical motifs as well as our understanding of the TOC34 GTP hydrolysis cycle. Furthermore, this study has advanced methodologies for transient expression and *in* vivo localization imaging using *Pisum sativum*. Together, our results have made a significant step forward in our understanding of the plastid protein import cycle.

TABLE OF CONTENTS

LIST OF FIGURES

ABBREVIATIONS USED

aa:	amino acid
AUC:	Analytical ultracentrifugation
BME:	β-mercaptoethanol
ER:	Endoplasmic reticulum
Fd-tp:	TP of ferredoxin
GAP:	GTPase activating protein
GAD:	GTPase activated by dimerization
GEF:	Guanine exchange factor
GB:	Grinding buffer
IB:	Import buffer
IC_{50}:	Inhibitor concentration at half maximum
IM:	Chloroplast inner membrane
IMS:	Chloroplast inner membrane space
K_d:	Equilibrium dissociation constant
MALDI-TOF MS:	Matrix-assisted laser desorption ionization-time of flight
mSSU:	Mature domain of prSSU
OM:	Chloroplast outer membrane
prSSU:	Precursor of the small subunit of RuBisCO
RuBisCO:	Ribulose 1,5-bis-phosphate carboxylase/oxygenase
SStp:	TP of the small subunit of RuBisCO
TIC	Translocon of the inner chloroplast membrane
TOC:	Translocon of the outer chloroplast membrane
TP:	Transit peptide
V_{max}:	Maximal velocity of an enzyme
YFP:	Yellow fluorescent protein

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

Disclosure:

Portions of this chapter have been published in a review paper: Chotewutmontri P.,* Holbrook. K* and Bruce, B. 2016. Plastid Protein Recognition and Import. International Review of Cellular and Molecular Biology. In press. (* denotes equal authorship).

1.1 Origin of Plastids

Robert C. Hooke first proposed the term “cell” 350 years ago when he observed a thin slice of cork under a simple microscope. Cells are now considered to be the smallest functional component of living organisms, and are functionally defined by their compartmentalization and various organelles. The membranes that define these organelles serve an important role in both intracellular communication as well as exchange of materials. For a plant cell, the plastid organelle is a key component in energy conversion.

Over century ago, plastids were proposed to be derived from cyanobacteria (Martin and Kowallik, 1999). It is now widely accepted based on phylogenetic analysis that the primary plastids originated from endosymbiosis of a cyanobacterial ancestor (Baum, 2013; Criscuolo and Gribaldo, 2011; Douglas, 1998; McFadden, 2014; McFadden and van Dooren, 2004). Interpretation of the fossil record and molecular clock analysis date the endosymbiosis event to over 1.2 billion years ago (Butterfield, 2000; Yoon et al., 2004). Cyanobacterial endosymbiosis provided the advantage of photosynthetic capability, and many cyanobacterial genes were transferred to the nucleus. The ability of plastids to import nuclear-encoded proteins enabled the plastid ancestry to transfer its genes to the nucleus without compromising metabolic capacity (Allen, 2003) (Figure 1-1). This process allowed endosymbiont gene copies to undergo pseudogenization and later loss, which resulted in a reduction in plastid genome size (Martin et al., 1993). Thus, most plastid proteins are synthesized in the cytosol and are post-translationally imported into their functional location inside the plastid. While plastid protein import was crucial in facilitating the endosymbiosis event in the past, it is indispensable to the function of the cells in *Arabidopsis* where over 2,500 proteins in *Arabidopsis* must gain access to plastids.

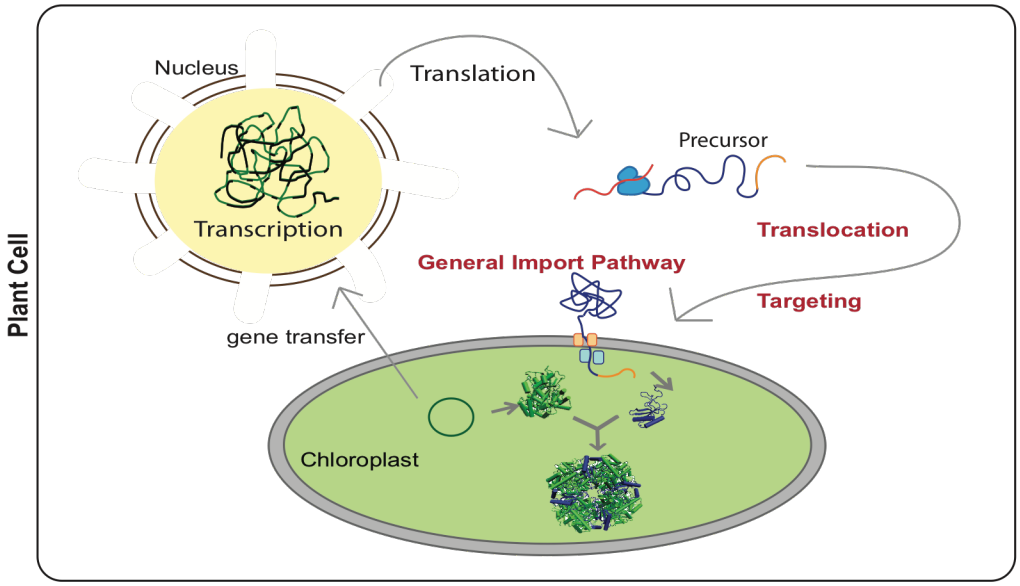


Figure 1-1 Plastid Protein Import

Cartoon schematic depicting the import of proteins into a plastid. Many proteins functional in the chloroplast are encoded by the nucleus and translated in the cytosol as precursor proteins with a short, N-terminal extension called the transit peptide. The transit peptide is required for the proper targeting and translocation of precursor proteins into their functional location in the plastid.

1.2 Plastid Protein Targeting Routes

Proteins targeted to plastids can be delivered post-translationally to six plastid locations: the outer envelope membrane (OM), the intermembrane space (IMS), the inner envelope membrane (IM), the stroma, the thylakoid membrane and the thylakoid lumen (Keegstra and Cline, 1999; Li and Chiu, 2010). Most of these processes involve protein translocation across the plastid membranes: the outer envelope, the inner envelope and the thylakoid membranes. In general, protein translocation across or insertion into membrane is mediated by oligomeric membrane complexes termed translocons (Walter and Lingappa, 1986). The majority of plastid-targeted proteins are synthesized as precursors containing the N-terminal targeting sequences called transit peptides (TPs) (Dobberstein et al., 1977; Kleffmann et al., 2004). Similar to the signal hypothesis (Blobel, 1980), TPs act as an intrinsic signal on the precursor protein that is recognized by the targeting receptors associated with the translocons. TPs direct translocation of precursor proteins across the double membranes of plastids via the translocon at the outer envelope of chloroplasts (TOC) and the translocon at the inner envelope of chloroplasts (TIC) in a process described as the general import pathway (Cline and Henry, 1996; Schnell et al., 1997; van 't Hof and de Kruijff, 1995a). After the precursor is translocated into the stroma, the TP is readily cleaved allowing the mature domain to fold into its native conformation or to be further targeted to the thylakoid (Richter and Lamppa, 2002a). So far, the TP has been utilized in targeting precursor proteins to all of the six locations of plastids and is considered to function in the general import pathway. This pathway recognizes diverse TP sequences from various functional groups of proteins. In summary, the general import pathway functions similar to a central transit hub where the majority of plastid-targeted proteins pass through before reaching their final locations.

1.2.1 The general import pathway

The general import pathway is a working model describing a TP-directed translocation of proteins into the stroma of plastids (Bruce, 2000a). Figure 1-2 illustrates the sequential steps in the pathway that reflect the processes that take place in the cytosol, at both the OM and IM, and in the stroma. We have not included steps that are involved in the subsequent targeting to the thylakoid nor do we address the targeting to the OM, IM or IMS. Figure 1-2 represents a somewhat selective model that has consolidated many

years of work, but emerging research continues to develop a more complete picture of the import system.

The OM of plastids has been shown to recruit cytosolic precursors using multiple pathways. Work performed using isolated chloroplasts shows that the precursors are able to directly bind to TOC on the surface of chloroplasts (Figure 1-2, step **1a**). The binding is reversible when lacking ATP/GTP and denoted as energy-independent binding (Jarvis, 2008a; Kouranov and Schnell, 1997; Perry and Keegstra, 1994). Addition of GTP also promotes binding (Inoue and Akita, 2008; Young et al., 1999). TP interaction with chloroplast lipids suggests that the precursor can directly bind to the lipid surface (Figure 1-2,step **1e**) before being transferred to the membrane receptors TOC159 and TOC34 (Figure 1-2, step **2f, 3a**) (Pilon et al., 1995a; Pinnaduwage and Bruce, 1996). Cytosolic factors were shown to interact and recruit precursors to the membrane receptors. Cytosolic Hsp90 captures precursor (Figure 1-2, step **1d**) and delivers it to the membrane receptor TOC64 (Figure 1-2, step **2d**) (Qbadou et al., 2006). Phosphorylated TPs interact with 14-3-3 and form the guidance complex with cytosolic Hsp70 (Figure 1-2, step **1c**) (May and Soll, 2000b).The guidance complex was proposed to dock to membrane receptors TOC159 (Figure 1-2, step **2c, 3a**) or TOC34 (Figure 1-2, step **2b**) (May and Soll, 2000b; Qbadou et al., 2006).Another pathway utilizes the cytosolic TOC159 pool to deliver precursor to the membrane (Figure 1-2, step **1b, 2a**) (Hiltbrunner et al., 2001; Smith et al., 2004). When low levels of ATP (<0.1 mM) are present, with or without GTP, the precursor engages in an irreversible binding state, which is referred to as the early import intermediate (Inoue and Akita, 2008; Kessler et al., 1994; Kouranov and Schnell, 1997; Olsen and Keegstra, 1992; Perry and Keegstra, 1994; Young et al., 1999). The switching from energy-independent to energy-dependent bindings seems to involve TOC75 (Paila et al., 2016). In the presence of 0.1 mM ATP at 4°C, about 110 amino acids are buried in the translocons (Figure 1-2, step **4**) (Akita and Inoue, 2009; Inoue and Akita, 2008). When the temperature is increased to 25°C, about 130 amino acids are buried (Figure 1-2, step **5**) (Akita and Inoue, 2009; Inoue and Akita, 2008). At this step, the precursor spans the double membrane via the TOC75 and TIC20/21 channels and TP interacts with TIC110 (Akita et al., 1997; Chen and Li, 2007b; Inoue and Akita, 2008; Nielsen et al., 1997) forming the contact site between the double membranes

Figure 1-2: The steps of the general import pathway

Cytosolic precursors can be targeted to plastids directly by interacting with TOC (1a), or by interacting with the lipids (1e) before transferring to membrane Toc159 (2f) and reaching Toc34 (3a). The precursors can also interact with cytosolic Toc159 (1b) before transferring to Toc34 (2a). TPs can interact with Hsp90 (1d) before being targeted to Toc64 (2d) and further transferred to Toc34 (3b). Phosphorylated TPs can interact with the guidance complex (1c) composed of 14-3-3 and Hsp70c in cytosol. The precursors from guidance complexes can be transferred to membrane receptors Toc159 (2c) or Toc34 (2b). The binding state of precursor to the chloroplast can be subdivided based on GTP/ATP level and temperature of the system (4-5). When the ATP level is greater than 1 mM, the translocation process is initiated by Hsp93/cpHsp70 (6). When the precursors emerge into the stroma, stromal processing peptidase (SPP) will cleave TP from precursor proteins releasing the mature domain (7). TP will be further degraded by PreP1/PreP2 peptidases (8). Adapted from Chotewutmontri, Holbrook and Bruce 2016.

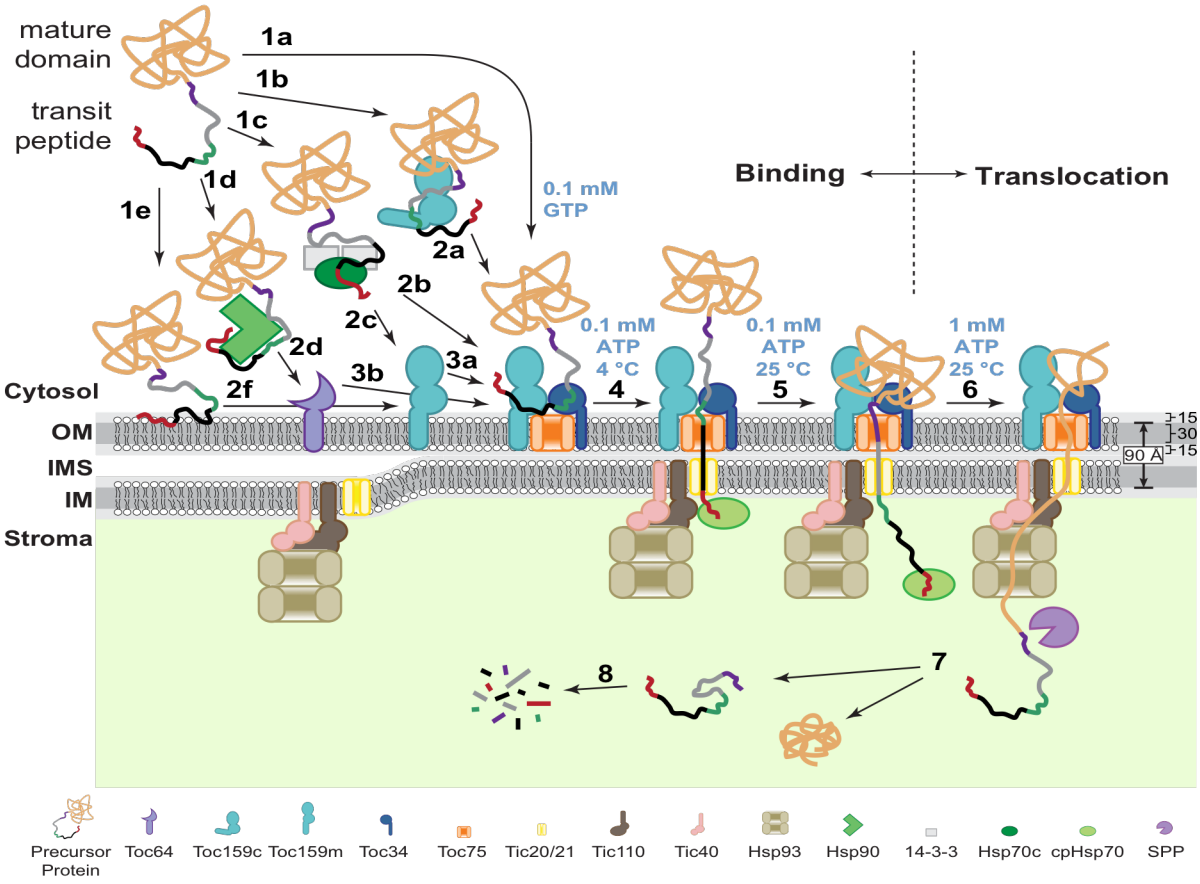


Figure 1-2 Continued

(Schnell and Blobel, 1993b). The translocation process is initiated by the ATPase chaperones Hsp93/cpHsp70 when the ATP level is above 1 mM (Figure 1-2, step **6**) (Shi and Theg, 2010b; Su and Li, 2010). When the precursor emerges into the stroma, stromal processing peptidase (SPP) will cleave TP from the precursor, releasing the mature domain (Figure 1-2, step **7**) (Richter and Lamppa, 1998). TP will be further degraded by presequence peptidases, PreP1/PreP2 (Figure 1-2, step **8**) (Glaser et al., 2006). The detailed mechanism of processes such as the specificity of TP recognition and threading of precursors into the stroma requires further study.

1.2.2 Roles of the Components of the General Import Pathway

1.2.2.1 Transit Peptides

Transit peptides (TPs) are N-terminal extensions that are cleaved during or after import. They were first identified forty years ago by Dobberstein and coworkers (Dobberstein, 1977) in *Chlamydomonas* for the small subunit of Rubisco and named transit peptides by Chua (1979) to be differentiated from the signal peptides. They are often encoded by the first exon and represent an addition of new coding information relative to the ancestral gene structure found in cyanobacteria (Kilian and Kroth, 2004; Tonkin et al., 2008).TPs are necessary and sufficient to facilitate protein import into plastids; the mature domain alone fails to be imported while addition of TP can direct the import of a non-plastid protein into a plastid (Bruce, 2001). Thus, TP contains information governing the import process. The length of TPs vary from 20-150 amino acids based on the position of the processing site (Balsera et al., 2009b). However, it has been established that short TPs cannot direct the import (Bionda et al., 2010). Import only occurs when short TPs were extended into their mature domains to reach at least 60-amino acids long (Bionda et al., 2010).

Many attempts have been made to identify the conserved motifs present amid diverse TPs using primary sequence alignment, but the conservation becomes greatly reduced with the increasing numbers of TPs analyzed (Karlin-Neumann and Tobin, 1986; von Heijne et al., 1989a). Amino acid composition and organization are also highly divergent (Bruce, 2000). Nevertheless, three regions were loosely defined: i) N-terminal domain of about 10 residues, lacking charged amino acids, ending with Pro/Gly and

preferably having Ala as the second residue, ii) central domain, lacking acidic amino acids but rich in hydroxylated amino acids, and iii) C-terminal domain, rich in Arg and possibly forming an amphiphilic β-strand (Bruce, 2001; von Heijne et al., 1989).

A few studies have determined NMR structures of TPs including those of TPs of RuBisCO activase (Krimm et al., 1999) and ferredoxin (Lancelin et al., 1994b) from *Chlamydomonas*, as well as the TP of ferredoxin from a higher plant *Silene latifolia* (Wienk et al., 1999). Even though TPs are found to be unstructured in aqueous solution, they were shown to adopt alpha-helical structures in membrane-mimetic environments inferring the possible formation and of an alpha-helix in TP recognition (Bruce, 2001a). Notably, the most stable alpha-helix of the TP present in higher plant ferredoxin contains a semi-conserved FGLK motif that was suggested to interact with the translocation apparatus (Schleiff et al., 2002b; Wienk et al., 2000). This is an example of the semi-conserved physicochemical motifs of TPs that are proposed to function as interacting domains determining organelle specificity.

1.2.2.2 Cytosolic factors

Precursors are proposed to maintain import-competency by interacting with cytosolic chaperones (Jackson-Constan et al., 2001). In fact, multiple Hsp70 binding sites have been shown in the majority of TPs (Ivey et al., 2000a; Rial et al., 2000; Zhang and Glaser, 2002). While cytosolic Hsp70 is essential for the *in vitro* import of a membrane protein, the precursor of the light harvesting chlorophyll a/b-binding protein (Waegemann et al., 1990), it is not necessary for the *in vitro* import of precursors of soluble proteins such as ferredoxin (prFD) (Pilon et al., 1990) and the small subunit of RuBisCO (prSSU) (Dabney-Smith et al., 1999).

Another biochemically identified cytosolic factor, the guidance complex composed of 14-3-3 and Hsp70 proteins, improves the efficiency of import (May and Soll, 2000b)14-3-3 recognizes the phosphorylated TPs (May and Soll, 2000b).This guidance complex was proposed to deliver the precursor to membrane receptors TOC64, TOC159 or TOC34 (May and Soll, 2000b). Later experiments showed that the complex associates with TOC34 but not with TOC64 (Qbadou et al., 2006). Furthermore, some precursors associate with another cytosolic chaperone, Hsp90. This chaperone delivers precursors to TOC64 through the interaction between tetratricopeptide repeats of TOC64 and Hsp90

(Qbadou et al., 2006). The precursor is subsequently transferred to TOC34 (Qbadou et al., 2006). In addition, TP receptor TOC159 (Ma et al., 1996; Perry and Keegstra, 1994) has been suggested to distribute both in soluble cytosolic and membrane-bound forms (Hiltbrunner et al., 2001) and it has been proposed that cytosolic TOC159 interacts with precursor proteins and shuttle them to the translocon.

1.2.2.3 Outer envelope lipids

The composition of lipids making up plastid envelope membranes resembles that of cyanobacteria (Joyard et al., 1991). These membranes contain high levels of glycerolipids, monogalactosyldiacylglycerol (MGDG), digalactosyldiacylglycerol (DGDG) and sulfolipid (SL), and low levels of phospholipids, phosphatidylcholine (PC), phosphatidylglycerol (PG) and phosphatidylinositol (PI). The composition of the lipids in plastids seems to be maintained over all plastid forms (Joyard et al., 1991). The presence of galactolipids, MGDG and DGDG make the outer membrane of the plastids unique among the cytosolically exposed membranes (Bruce, 1998).

The roles of lipids in the general import pathway have been analyzed mainly through the interactions with TPs (Pilon et al., 1995a; Pinnaduwage and Bruce, 1996; van 't Hof and de Kruijff, 1995b; van 't Hof et al., 1991; van 't Hof et al., 1993). TPs interact specifically with MGDG containing membranes (Pilon et al., 1995a; Pinnaduwage and Bruce, 1996), which has been proposed to be mediated by hydroxylated amino acids (Pilon et al., 1995a). A study using an *Arabidopsis dgd1* mutant containing less than 10% wild-type DGDG levels showed that the mutant chloroplasts import proteins to the stroma slower than the wild-type chloroplasts (Chen and Li, 1998). This import rate correlated with the reduction of the early import intermediate suggests that the increased proportion of MGDG in *dgd1* mutant envelope might trap the precursor at the energy-independent binding state (Chen and Li, 1998). Another study using an *Arabidopsis* MGDG synthase mutant (*mgd1-1*) producing 42% wild-type MGDG levels showed that protein import was not affected, which may be due to the high levels of remaining MGDG (Aronsson, 2008).

How TP-lipid interactions function in the import process remains unclear. However, the ability of TPs to reorient MGDG containing bilayers has been hypothesized to be involved (Bruce, 1998a; Chupin et al., 1994). Another possibility is based on the observations that TPs only adopt secondary structures upon binding to lipids. It was

proposed that the TP-chloroplast membrane interaction triggers the folding of a specific recognition element for protein import (Bruce, 1998).

1.2.2.4 TOC apparatus

1.2.2.4.1 Core complex

The TOC-TP recognition event is the first step in the general import pathway in chloroplasts. The TOC core complex is composed of 3 subunits, TOC159, TOC75 and TOC34, and forms a hetero-oligomeric complex ranging from 500 to 1,000 kDa (Chen and Li, 2007b; Kikuchi et al., 2006a; Schleiff et al., 2003a). (Jarvis, 2008a). Both TOC159 and TOC34 contain a GTPase domain (Kessler et al., 1994) while TOC75 is a beta-barrel protein (Hinnah et al., 1997). The subunits are named based on their predicted molecular weights in kDa (Schnell et al., 1997). The stoichiometry of these subunits is still debatable. Ratios of 4:4:1 and 6:6:2 for TOC34:TOC75:TOC159 have been proposed (Kikuchi et al., 2006a; Schleiff et al., 2003a). This complex size difference also indicates the possibility of dynamic complex forming higher order structure The size and components of the active TOC core complex are dynamic. Through the course of plastid differentiation, plastids can increase in size greater than 100-fold and proteins, lipids, and cofactors necessary for photosynthesis accumulate (Mullet, 1988). The chloroplast import apparatus must therefore be able to accommodate the increase in the expression of key photosynthetic proteins. The protein import capability of chloroplasts decreases over the course of development, and has been experimentally determined to decline as much as 20-fold in *in vitro* import assays (Dahlin and Cline, 1991). Thus, the core complex can be modified and its size depends on the metabolic and developmental state of the cell. It is clear that the core complex is dynamic and this modulation represents another level of regulation on the general import pathway.

1.2.2.4.2 TOC Complex Assembly

The exact sequence of events in TOC complex assembly remains unclear. However, studies have established that levels of TOC75 protein are significantly higher than both TOC34 and TOC159 at early stages of development during plastid biogenesis (Kouranov et al., 1998). Therefore, it is likely that “free” TOC75 is able to nucleate the assembly of new TOC complexes. The TOC75 channel could facilitate interaction

between the transmembrane helices of the TOC core complex proteins and the lipid bilayer (Hofmann and Theg, 2005b; Richardson et al., 2014). Analysis of atTOC75 revealed that POTRA domains recruit atTOC33 and atTOC159 and the deletions of POTRA domains altered TOC complex assembly (Paila et al., 2016), supporting the nucleation role of TOC75. Reconstitution of the complex from individual components has established that TOC159 insertion is dependent on both TOC75 and TOC34, suggesting that TOC159 is the final element of the core complex assembly (Wallas et al., 2003). Interactions between the G-domains of the TOC34 and TOC159 receptors likely maintain the stability and stoichiometry of the complex. This assembly mechanism would suggest that the receptor integration determines the rate of formation of new TOC complexes (Richardson et al., 2014). However, TOC75 is the common element in TOC complexes. Although the measured size of TOC complexes are heterogeneous, most complexes are found at over ~800kDa, suggesting that the core complex contains two molecules of TOC159 and six molecules each of TOC34 and TOC75 (Chen and Li, 2007a; Kikuchi et al., 2006b; Schleiff et al., 2003a). Excess TOC75 could provide a pool of nucleating protein for rapid assembly of distinct complexes, with varying integration of TOC34/TOC159 family members. Additional analysis of the targeting of the individual TOC core components is required to fully understand the factors that control complex assembly.

1.2.2.5 TIC apparatus

The TIC complex was identified to be composed of TIC20, TIC21, TIC22, TIC32, TIC40, TIC55, TIC62 and TIC110 (Bedard and Jarvis, 2005; Kessler and Schnell, 2006; Smith, 2006; Soll and Schleiff, 2004). Their functions in protein import range from facilitating precursor translocation through IMS, TOC-TIC formation, TIC channel formation, stromal motor function, and regulation of import. Many TIC subunits have been hypothesized to form the translocation channel of the inner envelope. However, the first proposed TIC channel was identified via electrophysiological measurement, called protein import related anion channel (PIRAC) (van den Wijngaard and Vredenberg, 1999). PIRAC opening is regulated by TP (Dabney-Smith et al., 1999; van den Wijngaard et al., 1999) and it was found to associate with TIC110 (van den Wijngaard and Vredenberg, 1999). Another electrophysiological experiment identified TP regulated cation channel

TIC110 as a TIC channel (Heins et al., 2002). This result excludes PIRAC as functioning directly as the TIC channel since all other protein translocation channels are cation selective (Heins et al., 2002). It was later found that TIC110 contains 6 transmembrane helices and the opening was regulated by Ca^{2+} (Balsera et al., 2009a). TIC110 is also essential for TIC complex formation (Inaba et al., 2005). The other potential TIC channels are TIC20 and TIC21, which are related to the pore-forming subunits of the mitochondria translocon in the inner envelope (Reumann and Keegstra, 1999; Teng et al., 2006). The *attic20* RNAi knockdown and *attic21* knockout plants showed defective translocation across inner membranes (Chen et al., 2002; Teng et al., 2006). While TIC20 is expressed mainly in early development, TIC21 is expressed at a later stage indicating a shared function of these proteins (Teng et al., 2006). Components of the TIC system still require additional research to fully understand their role in the import cycle.

1.2.2.6 Stromal motors

It is well established that internal ATP hydrolysis is the driving force for plastid protein translocation (Theg et al., 1989a). Since the identification of Hsp70 family of proteins involved in protein import into endoplasmic reticulum (ER) and mitochondria (Chirico et al., 1988; Deshaies et al., 1988; Murakami et al., 1988; Zimmermann et al., 1988), ATPase chaperones have been proposed to drive the translocation of precursor proteins into plastids (Keegstra and Cline, 1999; Marshall et al., 1990). Two classes of chaperones, Hsp70 and Hsp100, have now been estabalished to associate with and facilitate the protein translocation.

Three Hsp70 proteins associated with chloroplasts were first identified from pea OM and stroma (Marshall et al., 1990). Later, a Hsp70 was identified in the pea chloroplast OM as an import intermediate associated protein (IAP) (Schnell et al., 1994). This IAP can be detected by antibodies against mammalian cytosolic Hsp70, mAb SPA-820 (Schnell et al., 1994). Because the Hsp70 IAP was protected from an OM impermeable protease thermolysin, it was proposed to localize in the IMS and interact with incoming precursor proteins translocated through TOC75 (Schnell et al., 1994). The gene corresponding to the Hsp70 IAP has not yet been identified. A later study, however, found that the protein recognized by mAb SPA-820 is located in the stroma (Ratnayake et al., 2008). The other outer membrane Hsp70s from pea and spinach, the thermolysin-

sensitive Com70s, were shown to associate with various precursors (Ko et al., 1992; Kourtz and Ko, 1997). Com70 also has higher sequence similarity with mammalian cognate Hsp70s than *E. coli* Hsp70, DnaK (Ko et al., 1992). It was later suggested that Com70 is the cytosolic Hsp70-1 (Guy and Li, 1998; Li et al., 1994). A stromal Hsp70 was later shown to associate with the translocation complexes (Nielsen et al., 1997).

In *Arabidopsis*, the only 2 Hsp70s predicted to harbor TP were shown to localize in the stroma (Ratnayake et al., 2008; Su and Li, 2008; Sung et al., 2001). These stromal Hsp70s are closely related to the bacterial Hsp70, DnaK (Ratnayake et al., 2008) and were named atcpHsc70-1 and atcpHsc70-2 (Su and Li, 2008). These cpHsc70s were shown to be directly involved in protein import. The *atcphsc70-1 atcphsc70-2* double knockout is lethal and the single knockout plants showed reduced protein import efficiencies (Su and Li, 2010). Biochemical analysis also found atcpHsc70s stably associate with the translocons and the precursors (Su and Li, 2010). The stromal Hsp70, ppHsp70-2 involved in protein import was concurrently identified in moss *Physcomitrella patens* (Shi and Theg, 2010b).

The stromal Hsp93 (ClpC), a member of the Hsp100 chaperone family, was first identified to be stably associated with the translocons (Nielsen et al., 1997). Two homologs were found in *Arabidopsis*, atHsp93-V and atHsp93-III (Kovacheva et al., 2005). Crosslinking data suggested that Hsp93, TIC110 and TIC40 function in concert during protein import (Chou et al., 2003), consistent with genetic analysis results (Kovacheva et al., 2005). The molecular interactions between these proteins have been studied. The binding of TP to TIC110 triggers the association of TIC110 with TIC40, which in turn induces the release of TP from TIC110 (Chou et al., 2006). TIC40 was also shown to stimulate Hsp93 ATP hydrolysis. Because it has lower affinity to ADP-bound Hsp93, TIC40 was suggested to act as an ATPase activating protein of Hsp93 (Chou et al., 2006). Nonetheless, a mutant of Hsp93 that obviates its ability to interact with Clp proteolytic core demonstrates the role of Hsp93 in quality control (Flores-Pérez et al., 2016).

Both Hsp70 and Hsp100 systems are essential for viability of plants (Shi and Theg, 2011). Knockout of each system are lethal such as those observed in *atcphsc70-1 atcphsc70-2* double knockout (Su and Li, 2008), *athsp93-V athsp93-III-2* double knockout (Kovacheva et al., 2007), and *pphsp70-2* knockout (Shi and Theg, 2010b). However,

these results may reflect not only their roles in protein import but also other roles in plant development (Constan et al., 2004; Lee et al., 2007; Su and Li, 2008). In minimally invasive cases, the knockout lines of a single homolog of each system, such as *atcphsp70-1, atcphsp70-2,* and *athsp93-V* single mutants, import efficiencies dropped to about 40-60% (Su and Li, 2010) indicating important roles in protein import of each system. When the functions of both systems were reduced by knocking out a single homolog from each system such as the *atcphsp70-1 athsp93-V* double knockout plant, import was further reduced to 30% (Su and Li, 2010). This additional reduction compared to the effect of each single mutant, together with the coimmunoprecipitation of Hsp70 and Hsp93 suggested that both systems act in concert in the same complex (Shi and Theg, 2011).

Interesting, quantitative analysis to determine the energetics of preprotein import have shown it to be an energy intensive process. Using two different precursors prepared by three distinct techniques, Shi and Theg have shown that the import of a precursor protein into chloroplasts is accompanied by the hydrolysis of ~650 ATP molecules. This translates to a ΔG protein transport of some 27,300 kJ/mol protein imported. When the cost of preprotein import is expanded to include the complete biogenesis of the organelle, they estimate that protein import across the plastid envelope membranes consumes ~0.6% of the total light-saturated energy output of the organelle (Shi and Theg, 2013).

1.2.2.7 Peptidases

The processing of prSSU was discovered nearly 40 years ago (Dobberstein et al., 1977). This finding led to the identification of two SPPs from pea (Oblong and Lamppa, 1992; VanderVere et al., 1995) and one in *Arabidopsis* (Richter and Lamppa, 1998; Zhong et al., 2003). SPP is classified as a member of the metalloprotease M16B subfamily with a zinc-binding motif (Aleshin et al., 2009; Richter and Lamppa, 1998; Richter et al., 2005; VanderVere et al., 1995). It was shown that SPP is the general processing enzyme of plastid protein import by its ability to proteolyze various precursors (Richter and Lamppa, 1998). SPP specifically binds to the C-terminal 12 residues of TP of prFD (Richter and Lamppa, 1999, 2002b, 2003). *In vitro*, the TP directs precursor binding to SPP (Richter and Lamppa, 1999). The mature domain is released immediately after cleavage of TP while TP remains attached (Richter and Lamppa, 1999). TP is further

fragmented before release from SPP (Richter and Lamppa, 1999). The SPP recognition sites at the C-terminus of TP seem to share a weak binding motif (Emanuelsson et al., 1999; Gavel and von Heijne, 1990). Later, the physicochemical properties at specific residues were proposed to form a SPP binding motif (Richter and Lamppa, 2002b). In fact, the regions at the C-termini of TPs show a positive net charge and are conserved at the position -1 for basic residue and positions -4, -3, -2 and +1 for uncharged residues relative to the SPP cleavage sites (Richter and Lamppa, 2002b).

Free cleaved TPs in the stroma are potentially harmful (Glaser et al., 2006) and are degraded by presequence peptidases (PrePs) (Bhushan et al., 2005). PrePs were originally identified from the mitochondrial matrix as mitochondrial presequence degradation enzymes (Stahl et al., 2002). They are classified as members of the metalloprotease M16C subfamily (Glaser et al., 2006). *Arabidopsis* has 2 PrePs, atPreP1 and atPreP2. Both proteins were found to have dual localizations in both chloroplasts and mitochondria (Bhushan et al., 2003; Bhushan et al., 2005; Moberg et al., 2003; Stahl et al., 2002). In addition, these proteins have different tissue-specific expression patterns; atPreP1 is expressed highly in flowers and can be detected in sliliques while atPreP2 is expressed in leaves, shoots and roots (Bhushan et al., 2005).

1.2.3 Non-canonical trafficking

Although the general import pathway directs the transport of thousands of plastid targeted proteins, non-canonical pathways are critical for organelle function. The chloroplast proteomic data suggests that around 20% of chloroplast proteins lack any predictable targeting sequence (Kleffmann et al., 2004) and around 1-8% of chloroplast targeted precursors contain an ER signal peptide (Kleffmann et al., 2004; Zybailov et al., 2008). Thus, a small fraction of protein lacking TPs are still able to target to the plastids. In recent years, several alternate targeting pathways have been reported (Figure 1-3). Non-canonical transport pathways include vesicle mediated trafficking, import of tail-anchored/signal anchored membrane proteins, and import of proteins with dual targeting signals to the mitochondria and plastid. In this section we summarize the current understanding of non-canonical transport, although many aspects of these pathways remain unclear and the focus of ongoing research. Developing a clearer picture of the

integration of non-canonical transport with the general import pathway will provide insight into plastid biogenesis and organelle maintenance.

1.2.3.1 Vesicle mediated trafficking

Organelles cannot synthesize every molecule they require, so a variety of metabolites, lipids and proteins must be exchanged via the secretory pathway, which is well studied between the ER, Golgi and plasma membrane. Accumulating evidence suggests that both a cytosol to plastid and intra-plastid vesicle delivery system exist for both lipids and proteins (Karim and Aronsson, 2014; Khan et al., 2013; Lu, 2016). In agreement, *Arabidopsis* genome analysis, bioinformatics, and web-based localization prediction tools have confirmed that there are plastid localized proteins with high sequence similarity to the components of the cytosolic
vesicular system (Khan et al., 2013). Analyses have identified eight COPII related proteins, as well as proteins related to SNAREs and RABs (Andersson and Sandelius, 2004; Brandizzi, 2011). Furthermore, active GTPase homologs to SAR1 and dynamin were identified (Khan et al., 2013). Thus, intra-plastid vesicular trafficking appears to use a system analogous to COPII-mediated transport. However, even though putative components of the intra-plastid vesicular trafficking mechanism have been identified, substantial research is required to understand their specific roles in transport.

The evolutionary basis for plastid vesicle trafficking has been the focus of several studies. Interestingly, several of the proteins implicated in plastid vesicular trafficking are conserved in cyanobacteria (Keller and Schneider, 2013; Khan et al., 2013), implying that vesicle trafficking may have developed before endosymbiosis. However, there are also clear similarities with the cytosolic system, suggesting possible transfer of secretory pathway components to the plastid via divergent evolution (Vothknecht and Soll, 2005). Analysis of organisms from diverse lineages suggests a late evolutionary development of plastid vesicular trafficking (Westphal et al., 2003). However, new phylogenetic analyses of *Arabidopsis thaliana* and *Oryza sativa* suggests that vesicular trafficked proteins are of eukaryotic (not cyanobacterial) origin (Gagat et al., 2013). The early eukaryotic host-derived proteins were already targeted to the endomembrane system by their signal peptide and pre-adapted to be targeted to plastids via the same system. Furthermore, endomembrane transport was the only viable plastid import route for several eukaryotic

proteins that require post-translational modifications or dual targeting to the plastid and cell wall. The phylogenetic results are consistent with a later development of endomembrane trafficking to plastids (Gagat et al., 2013). It is likely that the system developed in response to selective pressure to maintain a complex thylakoid membrane (Karim and Aronsson, 2014), which is spatially separated from the chloroplast envelope by the aqueous stroma in higher plants.

The lipids that comprise the thylakoid are synthesized through a pathway that involves exchange of lipid precursors from the ER to the IM; the final lipid products are produced in the IM but must reach their destination in the thylakoid (Douce and Joyard, 1990). Although there are several hypotheses for lipid transfer, both ultrastructural and biochemical evidence supports a vesicular mediated trafficking system for lipid movement to the thylakoid (Karim and Aronsson, 2014). Although vesicular lipid movement has been characterized, protein targeting through vesicles is less defined. Furthermore, it remains unclear if vesicles only play a role in intra-plastid localization of proteins, or if proteins can be targeted from the cytosol to the plastid via vesicles. Two such proteins that utilize vesicle trafficking from the cytosol contain signal peptide for ER targeting, carbonic anhydrase 1 (Villarejo et al., 2005) and nucleotide pyrophosphatase/ phosphordiesterase 1 (Nanjo et al., 2006), were shown to localize to plastids. Additionally, a recent study searched for nuclear encoded proteins that may be trafficked to the thylakoid via vesicles. Bioinformatic analysis identified homologs to thylakoid localized proteins with a COPII selection motif and found 21 transmembrane and 12 soluble candidates (Khan et al., 2013). About 45% of these candidates are linked to photosynthesis, implying that a subset of nuclear encoded proteins are trafficked to the thylakoid via vesicles (Khan et al., 2013). Furthermore, RNAi knockdown of a candidate SAR1 homolog GTPase (*cspar1*) resulted in downregulation of several nuclear encoded photosynthetic proteins, suggesting they are potentially trafficked via vesicles (Bang et al., 2012). Thus, both bioinformatics and experimental data support a role for vesicles in both intra-plastid delivery and cytosol to plastid mediated protein transport. Developing a better understanding of this system could have important implications in plastid biogenesis, thylakoid maintenance and assembly of the photosynthetic apparatus (Lu, 2016).

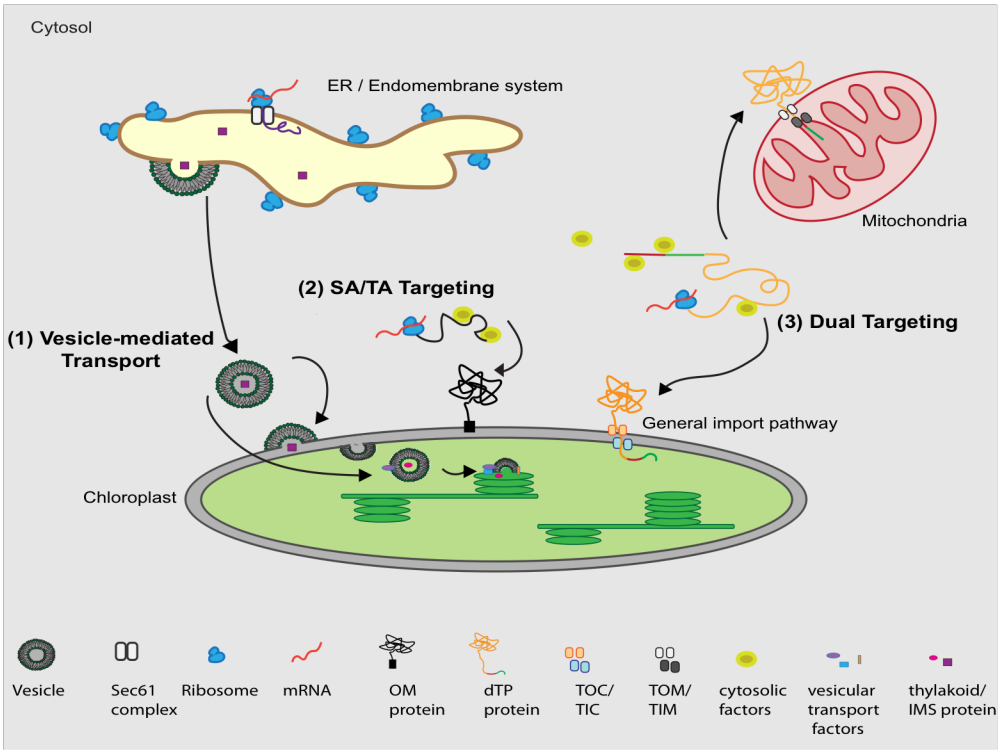


Figure 1-3: Non-canonical trafficking pathways into the plastid

Schematic depicts noncanonical plastid trafficking pathways: (1) Vesicle trafficking of a protein from the ER/endomembrane system to the plastid; vesicle trafficking of protein/lipids from the IM to the thylakoid (2) SA/TA protein cytosolic translation and insertion into the plastid OM (3) dual targeting of a protein to the mitochondria and plastid through the TOM/TIM (mitochondria) and TOC/TIC complexes.

1.2.3.2 Signal anchored and tail anchored proteins

Outer membrane proteins play a key role in intracellular communication, including exchange of metabolites and lipid synthesis. At least six pathways mediate protein targeting to the outer envelope of plastids (Hofmann and Theg, 2005a; Jarvis, 2008a; Richardson et al., 2014). Furthermore, correct localization of plastid outer membrane proteins is critical for the function of the general import pathway. Outer membrane proteins are a group of diverse proteins that can be divided into alpha-helical transmembrane domain (TMD) containing proteins and beta barrels consisting of multiple transmembrane beta strands (Inoue, 2015). TMD containing proteins can be further classified based on the number and location(s) of their TMD sequence (N-terminal, middle or C-terminal) (Lee et al., 2014). These proteins are nuclear encoded and synthesized on cytosolic ribosomes; they have distinct, specific requirements for post-translationally targeting to their functional location. In plastids both signal anchored (SA) and tail anchored (TA) proteins play an important role in outer membrane communication. Although both SA and TA proteins have hydrophobic TMD domains that direct insertion into the plastid OM, the secondary signals that mediate organelle specificity are not well understood. Physicochemical domains near the TMD, such as positively charged residues, seem to play a role in trafficking (Lee et al., 2014). Cytosolic factors likely play a role in the post-translational insertion of SA/TA proteins. Some OM proteins also contain a specialized TP sequence. For example, TOC75 utilizes a bipartite targeting sequence composed of stromal TP and envelope targeting signal (Inoue and Keegstra, 2003; Tranel and Keegstra, 1996). The insertions of OEP14, TOC159 and TOC75 also require some components of the general import pathway. Other outer envelope targeting pathways have been shown but are not well characterized (Hofmann and Theg, 2005a).

The most studied pathway for OM proteins requires an N-terminal TMD targeting signal anchor, which is found in proteins such as the outer envelope protein of 14 kDa (OEP14) and the TOC subunit of 64 kDa (TOC64) (Hofmann and Theg, 2005b; Lee et al., 2001). SA proteins are a class of outer envelope proteins with a single TMD that acts as both the localization signal and membrane anchor (Rapoport, 2007). SA proteins lack a cleavable TP; a moderately hydrophobic TMD acts as the primary targeting signal. Another required motif for targeting of SA proteins appears to be a C-terminal positively

charged domain (Rapaport, 2003). However, distinguishing the similar localization signals for plastids versus mitochondria has proven challenging (Lee et al., 2014; Lee et al., 2011).

Another pathway involves a C-terminal “tail anchor” TMD targeting sequence such as those present in TOC34 and TOC159 (Smith et al., 2002; Tsai et al., 1999). TA proteins contain a single C-terminal TMD and basic residues in a C-terminal sequence. Similar to SA proteins, TA proteins lack a cleavable TP. Plastid TA sequences do not share any obvious conserved sequence motifs, but they have a significantly higher hydrophobicity index versus mitochondrial TA proteins (Lee et al., 2014).

Only a few plastid TA proteins have been studied in depth, including TOC complex receptors TOC33 and TOC34. The C-terminal TMD domain of these proteins is necessary but not sufficient for OM localization; the GTPase domain of these proteins is required as well (Dhanoa et al., 2010). However, the exact targeting mechanism is unclear, because a truncated form of the TOC159 GTPase domain lacking a TA domain still correctly localizes to the OM (Smith et al., 2002). Thus it appears that the TA sequence is not always sufficient and other sequence context is necessary for correct localization. The mechanisms that regulate SA/TA protein organelle specificity and insertion will be key to understanding organelle biogenesis and plant development in depth.

1.2.4 Dual Targeting to Plastids and Mitochondria

Analogous to plastids, mitochondria import more than 1,000 proteins that are nuclear encoded and synthesized in the cytosol as precursor proteins (Murcha et al., 2014). The mitochondrial protein import complex is a set of multi-subunit protein complexes that recognize and transport mitochondrial proteins to their functional location (Schulz et al., 2015). The majority of mitochondrial proteins destined for the mitochondrial matrix contain a cleavable N-terminal presequence that is necessary and sufficient for import (Murcha et al., 2014; Schulz et al., 2015). Mitochondrial presequence amino acid composition is remarkably similar to chloroplast TPs. Mitochondrial presequences exhibit a high concentration of hydrophobic and positively charged residues, with an abundance of both proline and glycine, and an absence of negative charge (Bhushan et al., 2006; Huang et al., 2009a). Comparison of TPs and presequences suggests that specificity is likely determined by precise structural and physicochemical motifs that are required for explicit

recognition and translocation (Bhushan et al., 2006; Huang et al., 2009b). However, a subset of proteins function in both mitochondria and chloroplasts. Mechanisms for dual import must balance organelle selectivity but retain the ability to target either organelle.

Organelle targeting is normally highly specific. Interestingly, in *in vitro* studies, mistargeting of chloroplast proteins into mitochondria has been reported, but mislocalization of mitochondria proteins into chloroplasts has never been observed (Bhushan et al., 2006; Cleary et al., 2002; Huang et al., 2009b; Lister et al., 2001; Murcha et al., 2014). The lack of *in vivo* evidence for chloroplast to mitochondria mistargeting implies that there are several factors that provide organelle specificity. Cytosolic factors have been implicated in providing an extra level of targeting control (Rudhe et al., 2002). The OM lipids of mitochondria and chloroplasts also have a proposed role in protein recognition and organelle specificity. Additionally, both mitochondrial presequences and chloroplast TPs contain intrinsically disordered regions in aqueous medium that form secondary structure in membrane mimetic environments, which may aid in organelle selection. The mature protein itself appears to contain signals that assist correct organelle localization, possibly through interaction with the protein import machinery on the outer organelle envelope (Dabney-Smith, 1999). Furthermore, mitochondrial presequences appear to contain a positively charged chloroplast import avoidance signal near the N-terminus (Ge et al., 2014). Thus, it is likely that specificity is driven by a combination of several physicochemical factors. However, exactly how a dual targeting signal can be recognized by distinct receptors is still unclear.

Dual targeted transit sequences have been divided into two major classes: twin presequences or an ambiguous dual targeting peptide (dTP) (Murcha et al., 2014; Peeters and Small, 2001; Silva-Filho, 2003). Twin presequences comprise two specific targeting domains that are positioned in tandem. Protein expression produces two proteins with alternate targeting signals because of two alternate in frame translation initiation codons (Danpure, 1995; Silva-Filho, 2003). Alternatively, dTPs involve a single ambiguous targeting sequence. Most dTPs contain overlapping targeting signals for both chloroplast and mitochondrial import (Carrie and Small, 2013). Thus, the dTP contains weak targeting signals for both organelles, and subtle sequence and physicochemical properties are essential for the fidelity of targeting. However, how a protein containing

weak targeting signals for both organelles “selects” an organelle remains unclear. This is likely regulated by both the metabolic state and developmental stage of the cell. dTP containing proteins include proteins involved in translation, DNA synthesis and maintenance, cellular protein folding, turnover and maintenance, and energy production (Carrie et al., 2009; Carrie and Small, 2013; Carrie and Whelan, 2013). It is not surprising that many of these proteins are trafficked to both organelles because similar processes are required for the function of both chloroplasts and mitochondria. Interestingly, bioinformatics and experimental analysis suggests that there is a larger proportion of dual targeted proteins than expected; over 500 *Arabidopsis* proteins are predicted to be dual targeted (Baudisch et al., 2014; Mitschke et al., 2009). Recent analysis suggests that dual targeted proteins are trafficked in the same pathways as organelle specific precursors (Baudisch et al., 2014; Langner et al., 2014). Therefore, understanding the signals that allow a dTP to be recognized by distinct translocons may also provide insight into canonical targeting pathways.

1.3.2 Expression control

Spatial and temporal expressions of the nuclear-encoded plastid precursor proteins under different internal and external conditions are well documented (Drea et al., 2001; Gesch et al., 2003; Harmer et al., 2000; Knight et al., 2002; Plumley and Schmidt, 1989; Vorst et al., 1990; Zhou et al., 2006). These regulations alter the cytosolic levels of precursors, which affect the rates of the precursor protein import. Because most nuclear-encoded plastid proteins utilize the general import pathway, changing the expression level of any of these proteins can also potentially affect the import rates of other proteins (Row and Gray, 2001a). Recent evidence indicates that some of the gene expression regulation of nuclear-encoded plastid proteins also involves retrograde signaling from the plastids (Estavillo et al., 2011; Gray et al., 2003; Kakizaki et al., 2009; Pesaresi et al., 2006). In addition to the precursor proteins, the expression of the translocon components are also highly regulated. While green tissues express higher levels of atTOC33, atTOC159, atTOC64-III, atTIC55, atTIC62 and atTIC40, non-green tissues express higher levels of atTOC34, atTOC132, atTOC120, atTIC20-I and atTIC20-IV (Gutensohn et al., 2000; Vojta et al., 2004). A transcription factor CIA2 was also shown to upregulate atTOC33 and atTOC75-III expressions in leaves (Sun et al., 2001; Sun et al., 2009). Thus, cells

control spatial and temporal protein import rates by altering precursor proteins levels and generate different combinations of translocon components.

1.3.3 Precursor specific import pathways

1.3.3.1 Photosynthetic and non-photosynthetic precursors

It was proposed that the multiple paralogs of the translocon subunits perform different functions (Jarvis et al., 1998). Currently, several lines of evidence support this hypothesis. Knockout mutant phenotype analysis and biochemical characterization found that atTOC33 associates with atTOC159 in the TOC complex functioning in the import of photosynthetic proteins while atTOC34, atTOC132, atTOC120 are found in the TOC complex that functions in the import of nonphotosynthetic proteins (Ivanova et al., 2004; Kubis et al., 2004; Smith et al., 2004). In spinach, two TOC34 isoforms were also identified (Voigt et al., 2005) suggesting that other plants also may utilize specialized TOC receptors. In addition, the A domains of TOC159 have been shown to function in precursor selectivity of atTOC159 and atTOC132 (Inoue et al., 2010). This selectivity is further shown to rely on the TP sequence of the precursors (Wan et al., 1996; Yan et al., 2006).

However, the element(s) of TP corresponding to precursor-class selection is still largely unknown (Jarvis, 2008a). One of the element discovered was a segment on the TP of *Arabidopsis* small subunit of RuBisCO (atSStp) from residue 41 to 49, which governs the TOC159-dependent pathway (Lee et al., 2009a). Another element was identified from microarray analysis of nuclear-encoded plastid protein genes in *ppi1* mutant, the *atTOC33* knockout plant (Vojta et al., 2004). Only the down-regulated genes were shown to contain positively charged amino acids at the C-terminal of TPs (−8 and −1 positions) suggesting this element is involved in atTOC34 recognition (Vojta et al., 2004).

The precursor-specific pathways seem to merge at the TIC complex where TIC components were found to associate with both photosynthetic and nonphotosynthetic proteins (Chen et al., 2002; Jarvis, 2008a; Kovacheva et al., 2005). Nevertheless, atTIC20-IV was suggested to function in the alternative import pathway for housekeeping proteins (Kikuchi et al., 2013).

1.3.3.2 Age-specific precursors

Recently, the age-dependent regulation of protein import has been discovered (Teng et al., 2012). Based on the optimal import rates, precursors can be classified into 3 different groups based on the age of chloroplasts: young chloroplast specific, old chloroplast specific, and age-independent. The import efficiency into different ages of chloroplasts was shown to depend on the sequence of the TP (Teng et al., 2012). Import competition assays also found that TPs competed better within their own groups, suggesting each group utilizes a specific pathway (Teng et al., 2012). The attempt to determine the age-specific signal of TPs identified two consecutive positively charged residues as a signal for the old chloroplast specific pathway (Teng et al., 2012). It is still unknown whether specific TOC receptor combinations participate in this recognition or if post-translational modification is involved in creating the age-dependent signal of TP (Teng et al., 2012). Although the physiological relevance of age-dependent import was shown by analyzing the precursor gene families, where each precursor contained a TP from a different age-selective group (Teng et al., 2012), it is unknown as to whether the aging of chloroplasts only depends on the age-selective import and/or differential expression of the precursors.

Nevertheless, the only reported components of translocons that differentially function at different ages are atTIC20 and atTIC21. Whereas atTIC20 function is important in the early development stage, TIC21 function becomes dominant in the mature stage (Li and Chiu, 2010; Teng et al., 2006).

1.3.4 Redox regulation

The redox regulation of plastid protein import was shown to occur at both TOC and TIC translocons (Balsera et al., 2010). Earlier studies found that Cys-modifying agents (Friedman and Keegstra, 1989; Row and Gray, 2001b) and disulfide reducing agents (Pilon et al., 1992b; Stengel et al., 2009), inhibit and stimulate protein import, respectively. Protein import in *Physcomitrella* and *Chlamydomonas* were also enhanced in the presence of reducing agents (Stengel et al., 2009). In addition, the oxidant $CuCl_2$ was found to inhibit protein import by inducing disulfide bridge formation between TOC34, TOC75 and TOC159 (Seedorf and Soll, 1995). Disulfide bridge dimerization of TOC34 with a single conserved Cys has also been shown both *in vitro* and *in organello* (Lee et

al., 2009b). These findings indicate the possibility of redox-dependent disulfide bridge regulation of protein import.

Another level of redox regulation involves TIC subunits. It was proposed that TIC components containing redox-related domains might be involved in regulation (Bedard and Jarvis, 2005). While both dehydrogenases TIC62 and TIC32 harbor NADPH-binding sites (Chigri et al., 2006; Stengel et al., 2008), TIC55 has a Rieske 2Fe-2S center (Caliebe et al., 1997). Additionally, TIC62 contains a binding site for ferredoxin-$NADP^+$ reductase (FNR) (Stengel et al., 2008). The ratios of stromal $NADP^+$/NADPH have been shown to regulate the movement of TIC62 between stroma and inner envelope, and the interaction of TIC62 with FNR (Stengel et al., 2008). In reducing condition, TIC62 accumulates in the stroma and has a higher affinity to FNR (Stengel et al., 2008). Another study showed that NADPH abolished TIC62 and TIC32 interaction with TIC110 (Chigri et al., 2006). Lastly, the stromal $NADP^+$/NADPH ratio has been linked to regulate chloroplast protein import of a subgroup of precursors where higher ratios stimulate import (Stengel et al., 2009). This result confirms the role of redox regulation in protein import. Further studies would be required to determine the exact mechanism controlling the redox regulation.

1.3.5 Phosphorylation regulation

Phosphorylation of TPs has also been shown to regulate protein import. Phosphorylation of multiple precursors have been observed (Chen et al., 2014; Lamberti et al., 2011b; Waegemann and Soll, 1996) together with the identification of the kinases (Martin et al., 2006). The cytosolic guidance complex is proposed to recognize phosphorylated precursors before delivering them to the translocons (May and Soll, 2000b). Although the phosphorylation consensus sequence (P/G (X) (K/R) X(S/T) X(S*T*) (May and Soll, 2000b) is not present in the sequence of all chloroplast targeted proteins, Ser/Thr residues are highly abundant in TPs. Furthermore, knockout of cytosolic kinases shows a clear effect in chloroplast differentiation (Lamberti et al., 2011b). A phosphorylation-dephosphorylation cycle has been suggested, where the incoming precursors are in phosphorylated forms and a translocon-associated phosphatase dephosphorylates the precursors to initiate translocation (Waegemann and Soll, 1996). However, the phosphatase has not yet been identified. Nevertheless, the mutants of three TPs lacking the phosphorylation sites were able to direct the import of GFP into the

plastids, which indicates that phosphorylation of TP is not required for plastid targeting (Nakrieko et al., 2004). Phosphorylation-dephosphorylation cycles may ensure that a subset of highly expressed proteins reach the plastid efficiently (Holbrook et al., 2016; Nakrieko et al., 2004)

1.3.6 Regulation by ubiquitin-proteasome system

The level of precursor proteins in cytosol has been shown to be regulated by the ubiquitin-proteasome system, specifically through the cytosolic Hsc70 and the C-terminus of Hsc70-interacting protein (CHIP) E3 ubiquitin ligase pathway (Lee et al., 2009c; Shen et al., 2007). Interestingly, other evidence also indicated that a putative C3HC4-type really interesting new gene (RING) E3 ubiquitin ligase SP1 interacts with all of the TOC components and initiates their degradation *via* proteasome (Ling et al., 2012). It was proposed that E3 ubiquitin ligase SP1 regulates the turnover of TOC components through ubiquination and degradation pathways. In combination with the differential expression of TOC components, this results in modulation of the composition of TOC components (Ling et al., 2012). SP1 mediated TOC composition change was also suggested to control the transition between plastid types (Ling et al., 2012). Clearly, regulation of both the complex components and precursor proteins work together to modulate import to the plastid. This is a highly complex system and additional analysis of regulation pathways are still required to fully understand how these pathways interplay.

1.4 Concluding Remarks

Significant research has focused on the mechanisms that control import of nuclear encoded precursor proteins. The general import pathway represents a central hub for recognition, regulation and membrane translocation of precursors; this system plays a key role in organelle function and homeostasis. However, it is clear that noncanonical pathways work in concert with the TOC/TIC system to respond to developmental and physiological responses in the cell. The focus of our current study is to develop a detailed understanding of TOC receptor function as well as the elements guiding TP recognition and import.

CHAPTER 2: EXPERIMENTAL PROCEDURES

2.1 Bioinformatic analysis of TPs

2.1.1 Alignment of SStp

The full-length protein sequences of the small subunit of RuBisCO from green plants (*Viridiplantae*) were retrieved from UniProt database (release 2014_05) by query search using key words “Ribulose bisphosphate carboxylase small chain”, “taxonomy:33090” and “fragment:no”. A total of 429 sequences were matched using this criteria. Sequences with 100% identity were deemed redundant and removed using CD-HIT (Li and Godzik, 2006). The remaining 384 sequences were aligned using MAFFT (Katoh et al., 2005) which was shown to outperform the other programs in aligning the benchmark short linear motif data sets (Perrodou et al., 2008). The aligned sequences were submitted to MaxAlign (Gouveia-Oliveira et al., 2007) to remove poorly aligned sequences and to reduce gap positions. Four additional poorly aligned sequences were removed via visual inspection which resulted in the final 328 sequences. These sequences were realigned with MAFFT. The consensus and logo plot of the alignment were generated using Jalview program (Waterhouse et al., 2009). Only TP sequences were shown, identified based on the cleavage site of the tobacco sequence (Mueller et al., 1983).

2.1.2 Analysis of *Arabidopsis* TPs

Analysis of the *Arabidopsis* genome found 912 highly confidently predicted (using multiple prediction algorithms) plastid precursors, of which, 327 proteins contain a similar length of transit peptide at ~50-60 a.a. and 231 precursors contain at least one ‘FGLK’ motif. We selected a subset of 24 proteins based on the presence of FGLK motif around residues 28-39. Using multiple online tools (TargetP, ChloroP 1.1, iPSORT, Predotar, PredSL, ProtComp, Protein Prowler) these precursors were then further analyzed to predict their subcellular location. The output from these prediction tools was integrated into a mathematical scoring algorithm to select the 7 most confidently predicted chloroplast proteins, as described in Chapter 5.

2.2 Polymerase Chain Reaction

2.2.1 Amplification of DNA inserts

Some DNA inserts used for cloning were produced via polymerase chain reaction (PCR) of the desired sequence with specific primers. The template/construct names and primer

sequences are listed in Appendix 1, 2 and 3. High efficiency polymerase (TaKara ExTaq or Phusion HotStart polymerase MasterMix) were used to amplify DNA inserts. 20 μL reactions contained 0.4 ng/μL DNA templates, 1X ExTaq Buffer, 0.03 U/μL polymerase, 0.2 mM each/dNTPs, and 0.2 μM of each primer (ExTaq), or 1X MasterMix and 0. 2 μM of each primer (Phusion). The PCR conditions were set with an initial denaturation of 2 min at 94°C (20 sec initial denaturation for Phusion), followed by 20-25 cycles of denaturation/annealing/extension (15 sec denaturation at 94°C (5 sec denaturation for Phusion), 15 sec at annealing temperature, and 1 min/1kb amplicon at 72°C (30 sec per kb for Phusion)), a step of extension for 10 min at 72°C and a holding step at 4°C. Annealing temperatures were chosen based on the average melting temperature of the primers.

2.3 Construction of vectors

2.3.1 Vectors based on pET21D

The soluble, Δtransmembrane fragment of psToc34 was expressed in-frame with a C-terminal 6XHistidine tag (Reddick et al., 2007b). Site directed mutagenesis was performed using TaKara high fidelity polymerase according to Quik-Change instructions provided by Stratagene. All PCR-generated constructs were sequenced to ensure fidelity of amplification. Primers are shown in Appendix 1.

2.3.2 Vectors based on pAN187

In vivo SStp-NT (RuBisCO) small subunit sequence derived from *N. tobacum*) or FDtp (sequence derived from *S. latifolia*) assay constructs were generated in *pAN187* vector (Nelson, 2007) with a C-terminal YFP tag as previously described (Chotewutmontri et al., 2012). Site directed mutagenesis was performed using TaKara high fidelity polymerase according to Quik-Change instructions provided by Stratagene. All PCR-generated constructs were sequenced to ensure fidelity of amplification using the 35S F primer. Mutagenesis primers are shown in Appendix 2.

Concatemerized *Arabidopsis thaliana* TP sequences were ordered from Epoch Life Sciences (Missouri City, Texas) (Chapter 5) were amplified using primers in Appendix 3 and inserted into the CloneJet cloning vector using the manufacturer's instructions

(ThermoFisher). Confirmed CloneJet constructs were digested with NHE1 and NCO1 and inserted into the pAN187 expression vector using NHE1 and NCO1 sites. New constructs were named according to the TP number (1-7)and Rule number (1-6) (Figure 5-4).
Dual tag constructs (H-S-SStp) were expressed in the pET30a vector in frame with an N-terminal 6XHis-S protein tag. PCR amplification using TaKara high fidelity polymerase and site directed mutagenesis was performed based on Quik-Change instructions provided by the manufacturer (Stratagene). All PCR-generated constructs were sequenced to ensure fidelity of amplification.

2.3.3 Vectors for *Agrobacterium* based transformation

WT or mutant SStp-NT derived from pAN187 construct were directionally cloned into the binary vector pFGC19. Constructs were verified using DNA sequencing. Constructs were maintained in GV3101 Agrobacterium using the appropriate antibiotics.

2.4 Expression and purification of proteins

2.4.1 psToc34 recombinant expression and purification

To study interaction of Toc34 with TPs, Toc34 was expressed in E.coli BL-21 RIPL as described previously (Reddick et al., 2007b). The truncated, soluble GTPase domain of *P. sativum* Toc34 was expressed in frame with a 6xHis tag in the pET21D construct. Cells were cultured in Luria-Bertani (LB: 1% tryptone, 0.5% yeast extract and 1% NaCl) at 37°C and shaken at 225rpm until $O.D._{600}$ reached 0.4-0.5. Cells were induced with 1 mM (final concentration) IPTG and protein was expressed 3 hrs at 37°C with shaking at 225 rpm. Cells were collected and lysed in Buffer A (20 mM Na-phosphate, pH8.0, 1 mM $MgCl_2$, 500 mM NaCl, 0.1% Triton X-100 , 1 mM PMSF, 1 μM leupeptin and 1 μM pepstatin) with a French press. Samples were sonicated (30s sonication/ 45s cooldown for 5 min of total process time) to fragment DNA. The lysate was centrifuged at 20,000 x *g* for 30 min to remove insoluble proteins. Supernatent was loaded on Ni-NTA resin (Gold Biotech) and extensively washed using Buffer A + 15 mM Imidazole. Bound proteins were stored at -20°C in 1X GBS buffer (20 mM Tricine-KOH, pH 7.65, 1 mM $MgCl_2$, 50 mM NaCl, 1 mM β-mercaptoethanol with 20% glycerol) until the day of an experiment. Proteins were thawed and eluted using Buffer B (20 mM Tricine-KOH, pH 7.65, 1 mM $MgCl_2$, 50 mM

NaCl + protease inhibitors) and assayed for protein content using BCA or Bradford protein assays.

2.4.2 Recombinant expression and purification of peptides using the IMPACT system

Expression of TPs was performed as described previously (Reddick et al., 2007b) using the IMPACT system (New England Biolabs). Peptides were tagged at the C-terminus with an inducible self-cleavage Intein protease fused to a chitin binding domain (CBD). After purification, only Pro-Gly was still tagged to the peptides. E.coli ER2566 (New England Biolabs) transformed with pTYB2 constructs were cultured in LB and shaken at 225 rpm until $O.D._{600}$ reached 0.3-0.4. Expression was induced with 1 mM IPTG and cultured overnight at 25°C. Cells were harvested and lysed using Buffer C (20 mM Na-phosphate, pH 8.0, 500 mM NaCl, 0.1% Triton X-100, 1 mM PMSF, 1 μM leupeptin and 1 μM pepstatin using a French press and sonication (45s sonication/30s cooldown with 5 min total process time) to fragment DNA. Lysate was centrifuged at 20,000 x *g* for 30 min. Supernatent was loaded onto chitin resin (New England Biolabs) and washed extensively using Buffer D (20 mM Na-phosphate pH 8.0, 500 mM NaCl, and 1 mM EDTA). Cleavage was induced by replacing the column buffer with Buffer E (1 mM Na-phosphate pH 7.5 and 50 mM β-mercaptoethanol) and incubating overnight at 4°C. Peptides were eluted with Buffer E without β-mercaptoethanol and lyophilized to remove β-mercaptoethanol. Lyophilized peptides were stored at -80°C until use.

2.5 Pea Plant growth

Dwarf pea (*P. sativum* – Green Arrow cultivar) were acquired from J.W. Jung Seed Co. and stored at 4°C. Seeds were imbibed overnight in cold running tap water with aeration before being planted in 35cm x 50 cm x 39 cm metal flat with coarse vermiculite (for use in chloroplast isolation experiments) or in plastic containers with vermiculite (*Agrobacterium tumefaciens* vacuum transformations). About 400 mL dry seeds were used per flat. Peas were grown in a growth chamber with illumination at ~100 $\mu E/m^2/sec$ on a 14hr light and 10hr dark cycle. Transient protein expression/plant transformation protocols are described in detail in the method development section of Chapter 6.

2.6 Chloroplast isolation

Chloroplasts were isolated from *P.sativum* plants as previously described (Bruce et al., 1994). Briefly, plant tissues were diced with a food processor. Plant tissues were kept in the dark and buffers were kept on ice or at 4°C throughout the procedure. Grinding Buffer (GB: 330 mM sorbitol, 1 mM $MgCl_2$, 1 mM $MnCl_2$, 2 mM EDTA, 0.1% BSA, 50 mM HEPES-KOH, pH 7.3) was added to the diced tissue before homogenization using a Polytron homogenizer. We used about 400 mL GB per flat of peas. The homogenate was filtered through 2 layers of Miracloth (Calbiochem) and two layers of cheesecloth. The filtrate was centrifuged at 3,500 x *g* for 7 min to collect a chloroplast pellet. Import buffer (IB: 300 mM sorbitol, 50 mM HEPES-KOH, pH 8.0) was added to the chloroplast pellet and a paintbrush was used to resuspend the chloroplasts. The resuspension was loaded on top of a continuous Percoll / IB(GE Healthcare) gradient. The loaded gradients were centrifuged at 5,800 x *g* for 15 min. Intact chloroplasts were collected from the lower dark green band using a 14-gauge stainless steel needle. To remove Percoll, the collected fraction was diluted in a 2-fold volume of IB and centrifuged at 3,500 x *g* for 10 min. The pelleted chloroplasts were resuspended in IB before loading on a second Percoll continuous gradient. After removal of Percoll, chlorophyll content was assayed and chloroplasts were used immediately in further experiments.

2.7 Chlorophyll measurement

Chlorophyll was extracted from isolated chloroplasts by adding 10 μL of chloroplast suspension into 990 μL of 80% acetone. The suspension was mixed by vortexing for 1 min and centrifuged at 21,000 x *g* for 1 min to remove insoluble material. The absorbance of chlorophyll pigment in 80% acetone was assayed at 663 and 645 nm. The total chlorophyll in mg/mL was calculated based on the equation described by Arnon (1949).

2.8 Kinetic analysis of psToc34

2.8.1 Phosphate release assay for GTP hydrolysis

Phosphate release analysis, scintillation counting, and data fitting to the Michaelis-Menton equation were performed as reported in (Chotewutmontri et al., 2012; Reddick et al., 2008a; Reddick et al., 2007b).

GTP hydrolysis assays were performed as described in (Reddick et al., 2007b). Briefly, in a flat-bottom 96 well microplate, reactions were performed in GBS, 10 nM [γ-^{32}P]GTP, and varying concentrations of cold GTP, with a final reaction volume of 100 μL. Purified TOC protein was added to the mixture to start the reaction. At the times indicated in the figures (see Chapter 3), a 12.5 μL aliquot was removed and added to 200 μL of 10% w/v activated charcoal in 50 mM HCl, 5 mM H_3PO_4. The mixture was thoroughly resuspended using vigorously pipetting with a multichannel pipet. The charcoal suspension was transferred to a 96-well 0.2 μm filter plate (Pall Corb, Ann Arbor MI), and vauum filtered into a receiving 96 well microplate. An 8 uL aliquot of filtrate was added to 150 μL of MicroScint 40 scintillation fluid (PerkinElmer Life Sciences) and mixed well. Counting was performed using a 96-well microplate scintillation counter (TopCount NXT Microplate Scintillation Counter (PerkinElmer Life Sciences). Counts/min data were graphed as counts/min versus time using GraphPad Prism.

2.8.2 Data analysis

As described in (Reddick et al., 2007b), the experiment was set up so that the reaction rate was linear over a 30 min time course. The Δcpm/min values were calculated from the slope of the line generated by counts/min versus time. Using the known specific activity of the isotope, the enzyme concentration, and the counting efficiency, the slope was transformed into a reaction velocity (GTP hydrolyzed/min/μmol of TOC protein). To determine the reaction parameters V_{max} and K_m values, the cold GTP concentration was varied. Once the rate values had been calculated for each GTP concentration, the rates (y axis) were graphed against the total GTP concentration (substrate)(x-axis). These values were fit using the Michaelis-Menton equation:

$$v = \frac{V_{max}[S]}{K_m + [S]}$$

2.9 MALDI-TOF

Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) was performed using a Microflex mass spectrometer (Bruker Daltonics) similar to a previously described protocol (Reddick et al., 2007). Mass spectra were acquired in

positive ion mode. The peaks were identified using flexAnalysis software (Bruker Daltonics) and analyzed withFindPept (Gattiker et al., 2002).

2.10 Quantitative analysis of *in vivo* protein import assays using fluorescent imaging

We calculated relative intensity ratio measurements to assess the amount of imported protein. Measurements of camera noise subtracted epifluorescence images were taken using ImageJ software (Abramoff, 2004) essentially as described in (Chotewutmontri et al., 2012). Briefly, the intensity per pixel values from different areas in the images were calculated from the summation of intensity signals in the area divided by the total number of pixels in the area. A circlular area was drawn to fit around individual plastids to measure the plastid intensity per pixel. The same circlular area was enlarged to threefold diameter, so that the area shared the same center, but integrated signals from the cytosol. The cytosol intensity was calculated from the ring area between the former and enlarged circles. A rectangular area outside the fluorescing cell in the same image was used to calculate the background intensity per pixel. The background intensity value was subtracted from the plastid and cytosol intensity per pixel values. For each plastid, a ratio between the background removed plastid and cytosol intensity per pixel values were calculated. The ratio of intensity of each cell is the average of all plastid ratio values.

CHAPTER 3: COUPLING OF MOLECULAR MODELING WITH *IN VITRO* ENZYMATIC ANALYSES TO ELUCIDATE HOW THE TOC GTPASE MECHANISM IS COUPLED TO THE PRE-PROTEIN IMPORT CYCLE

Abstract:

Most plastid-localized proteins are nuclear-encoded and post-translationally imported from the cytosol. The mechanism for the selective translocation of precursor proteins appears to involve the recognition of the transit peptide by the TOC GTPases, which act as receptors or gatekeepers. It is widely hypothesized that the primary targeting specificity is mediated through specific interactions between the transit peptide and the TOC receptors, TOC34 and TOC159. Although it is known that both TOC34 and TOC159 are GTPases and bind precursor proteins, little is understood about how their intrinsic GTPase activity is involved in the import process.

To aid in the elucidation of this complex process, we have developed highly sensitive GTP-binding and hydrolysis assays to characterize the catalytic mechanism of the TOC34 GTPase component. We have extended our analysis of TOC34 to integrate molecular dynamics simulations and QM/MM calculations. Molecular modeling has allowed for the design of site-specific mutations near the active site of TOC34 so that we can investigate its catalytic mechanism in detail. We are currently testing mutant proteins using the soluble, truncated GTPase domain expressed and purified in *E. coli*. Next, we will enlist *in vitro* kinetic analyses to further test the characteristics of these mutants. Our work is expected to make a major step forward in the understanding of the TOC GTPase catalytic mechanism as well as its importance in the preprotein import cycle.

3.1 Introduction

The mechanism and contribution of the TOC GTPase cycle is a complex question, and represents a “black box” in the field. Although high resolution structural data and detailed kinetic analysis is available (Aronsson et al., 2010; Lumme et al., 2014; Reddick et al., 2007b; Schleiff et al., 2002a; Sun et al., 2002a) understanding of the mechanism of TP recognition, the detailed catalytic mechanism of TOC34 GTP hydrolysis, and even the contribution of the TOC GTPase cycle are lacking. We have developed a simple model (Figure 3-1) to describe the possible function of the TOC receptors. In this model, the homo/heterodimerized TOC receptors represent a “closed” system. Once GTP nucleotide or TP is added to the system, the TOC receptors monomerize or change conformation to allow the precursor protein to be “loaded” into the TOC complex. The

precursor protein is able to access the TOC75 channel, and the protein can be translocated into the plastid stroma. Our current studies focus on developing the framework to test many aspects of this model, including the GTPase catalytic mechanism of TOC34, homodimerization of TOC34 and detailed interaction analysis with TPs.

We have developed highly sensitive GTP hydrolysis assays to characterize the catalytic mechanism of the TOC34 GTPase component in detail using a variety of active site mutants. Furthermore, we have extended analysis of TP interaction with the TOC34 receptor component using a combination of fluorescence resonance energy transfer (FRET) and chemical crosslinking. TOC34 has been well established to directly interact with TPs (Aronsson et al., 2010; Chotewutmontri et al., 2012; Lumme et al., 2014; Reddick et al., 2007b). However, despite a great deal of research focus, the location of TP-TOC34 interaction is still unclear. To analyze the TP-TOC34 interaction in detail, we developed methodology using an array of solvent exposed single Cys mutants and a combination of techniques including fluorescence resonance energy transfer (FRET) and homobifunctional chemical crosslinkers bis-maleimidohexane (BMH) and copper phenanthroline (CuP) (Figure 3-2). Using homobifunctional chemical crosslinking, we can stabilize the homodimeric vs. monomeric form of TOC34 for analysis using biochemical methods such as SDS-PAGE. BMH is a sulfhydryl reactive crosslinker with a 13 Å spacer arm, while CuP is a zero length oxidant. Chemical crosslinking is useful for high throughput testing of a variety of different mutants and nucleotide conditions. By combining FRET experiments with crosslinking we can start to develop a “map” of the TP interaction sites on TOC34. Our work is expected to make a major step forward in the understanding of the TOC GTPase catalytic mechanism as well as its importance in the preprotein import cycle.

The mechanism of FRET involves a donor fluorophore in an excited electronic state which can transfer its energy to a nearby acceptor florophore through long-range dipole-dipole interactions. Conceptually, the rate of energy transfer depends on the spectral overlap between the donor emission and acceptor absorption spectra (Figure 3-2A). The distance separating the donor and acceptor molecule will affect the resonance energy transfer rate and allow for interaction to be detected. We can take this a step further by using different acceptor molecules as shown in Figure 3-3. For example, signal

from Acceptor fluorochrome 590 can detect distance differences between 30Å and 50Å, allowing the differentiation between long distance interactions. Similarly, Acceptor fluorochrome 700 can detect distance differences between 10Å and 30Å, allowing the differentiation between short distance interactions. Thus, FRET based studies facilitates dynamic monitoring of the TOC34-TP interaction.

3.1.2 TOC GTPase activity

Many studies have focused on understanding TOC receptor function, and our current work aims to build on the groundwork laid by previous studies. TOC159 and TOC34 are thought of as the receptors or "gatekeepers" of the TOC translocon. Both of these proteins are well established to hydrolyze GTP and bind to precursor proteins (Reddick et al., 2007b). Although the TOC GTPase cycle has been studied in detail *in vitro*, less is understood about how the system works in concert to contribute to precursor import. TOC34 was crystallized nearly 15 years ago (Sun et al., 2002a). Structurally, the overall fold of the TOC34 monomer is divided into a globular GTP binding domain (G-domain) and a small C-terminal alpha helical part. TOC34 is structurally remarkably similar to the well-studied small GTPase p21RAS. Structural superposition of Ras-GppNp and TOC34-GDP reveals that TOC34 has 6 sequence inserts I1-I6 (Figure 3-4), which, except for 15, have implications in dimerization and GTPase activity (Pai et al., 1989; Sun et al., 2002a) (Figure 3-4). Despite this high-resolution structural data, many aspects of TOC34 function as well as the physiological function of GTP hydrolysis on protein import remain unclear. Homodimerization of TOC receptors are common; atTOC33, psTOC34 and psTOC159 have been shown to dimerize (Koenig et al., 2008a; Koenig et al., 2008b; Oreb et al., 2011; Reddick et al., 2007a; Sun et al., 2002b; Yeh et al., 2007). Heterodimerizations between atTOC33 and atTOC159 and between psTOC34 and psTOC159 were also observed (Bauer et al., 2002; Becker et al., 2004; Hiltbrunner et al., 2001; Rahim et al., 2009; Smith et al., 2002). Dimerization of atTOC33 was determined to be involved in the transition of precursor proteins from binding to translocation state of the import (Lee et al., 2009b). At least three factors have been found to affect dimerization *in vitro*. (i) Monomer-dimer equilibrium of atTOC33 and psTOC34 depends on their concentrations with the dissociation constants (K_d) around 400 µM and 50 µM, respectively (Koenig et al., 2008a; Oreb et al., 2011; Reddick et al.,

2007a). Based on the closed proximity of TOC34 in the core TOC complex (Schleiff et al., 2003a), TOC34 is likely to favor a dimer form in the complex (Lumme et al., 2014). (ii) Nucleotide loading state is another factor that affects dimerization. In both atTOC33 and psTOC34 cases, the GDP-bound forms produce more dimers than the GTP- bound forms (Koenig et al., 2008a; Lumme et al., 2014). The same effect was also reported in the heterodimerization between atTOC159 and atTOC33 (Smith et al., 2002). Surprisingly, the addition of a GTP transition state analog, aluminum fluoride, shifted the equilibrium of both atTOC33 and psTOC34 to become exclusively dimers (Koenig et al., 2008b). These findings and recent work using FRET interaction analysis indicate that the dimerization is likely triggered by GTP hydrolysis (Lumme et al., 2014). Furthermore, it can also be speculated that with the presumed active TOC receptor, the monomeric GTP-bound state hydrolyzes GTP to become dimer, which subsequently traps it in the dimeric GDP-bound inactive state. The dimeric conformation was shown to prevent nucleotide exchange (Oreb et al., 2011) even though TOC receptors have higher affinities to GTP than GDP (Jelic et al., 2003; Reddick et al., 2007a). In agreement, TOC34 has been classified as a member of the GAD GTPase family, a subclass of small GTPases that are activated by nucleotide dependent dimerization (Gasper, 2009; Lumme et al., 2014). Putative GAD members have a role in a wide variety of cellular processes, including membrane fusion and fission (atlastin) (Bian et al., 2011), vesicle trafficking (dynamin)(Chappie et al., 2010), and cytokinesis (septin) (Sirajuddin et al., 2009). GADs possess all the structural elements of the classic “switch” mechanism of small GTPases, but reciprocally activate their catalytic sites in the dimeric state, rendering classic GEFs and GAPs unnecessary for activity (Gasper, 2009). Lastly, (iii) the effect of TP on TOC receptor functions has been the subject of many studies. GTP-bound forms of atTOC33 (Gutensohn et al., 2000), psTOC34 (Jelic et al., 2002b; Schleiff et al., 2002b) and psTOC159 (Becker et al., 2004; Kouranov and Schnell, 1997) have higher affinities to TP than the GDP-bound forms. TP is well known to stimulate TOC receptor GTP hydrolysis (Becker et al., 2004; Jelic et al., 2003; Oreb et al., 2011; Reddick et al., 2007a). Two separate roles of TP in GTP hydrolysis have been reported. First, it was shown that TP stimulated GTP hydrolysis of psTOC34 while maintaining the same nucleotide exchange rate suggesting a role of TP as a GAP but not guanine nucleotide

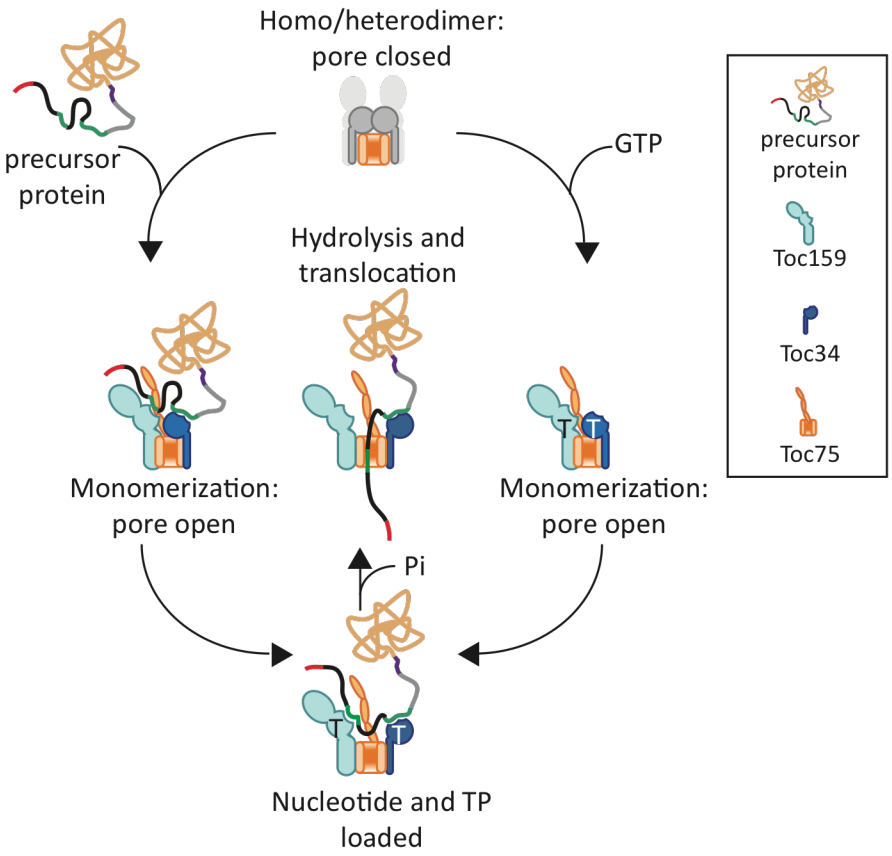


Figure 3-1: How does GTP hydrolysis contribute to the TOC import cycle?

Cartoon schematic showing predicted preprotein interaction with the TOC translocon. In this model, the receptor complex is monomerized (”opened”) through interaction with either GTP or TP. Subsequent GTP hydrolysis allows the precursor protein to pass through the beta-barrel TOC75. ATP hydrolysis by stromal chaperones pulls the precursor protein into the stroma.

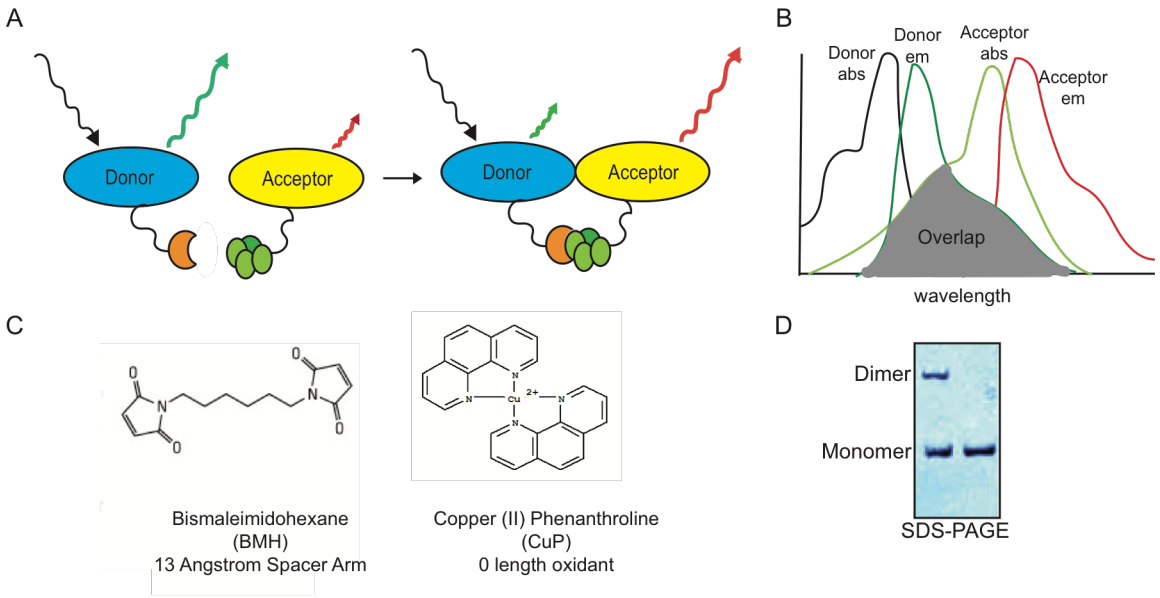


Figure 3-2: Investigating TP-TOC interaction with FRET, BMH and CuP

A. Model of FRET interaction. High energy wavelength fluorochrome attaches to cysteine on donor and low energy wavelength fluorochrome attaches to cysteine on acceptor. Donor is excited and placed with acceptor in fluorometer. FRET results in specific wavelength emission from acceptor. B. Example spectrum showing FRET expected results C. Chemical structures of crosslinkers BMH and CuP. D. Example SDS-PAGE showing expected results using chemical crosslinkers

A

$$E_{FRET} = 1/[1 + (r/R_0)^6]$$

R = Distance between residue and C215

R(0) is the characteristic distance where the FRET efficiency is 50 percent

B Calculated FRET Efficiency for Donor (495) and Acceptors (590 and 700)

DONOR 495	ACCEPTOR 590 R(0)=50	ACCEPTOR 700 R(0)=39
R=10Å	99.99%	99.97%
R=30Å	95.54%	82.84%
R=50Å	50.00%	18.38%

Figure 3-3: FRET efficiency measurements

A. FRET efficiency equation. B. Table shows calculated efficiency of the three chosen fluorochromes at different distances using the equation shown above. R(0) values were provided by the manufacturer (Attoc-Tec). Acceptor fluorochrome 590 shows a difference between 30Å and 50Å, allowing the differentiation between long distance interactions. Acceptor fluorochrome 700 shows a difference between 10Å and 30Å, allowing the differentiation between short distance interactions.

exchange factors (GEF) (Reddick et al., 2008b; Reddick et al., 2007a) since GAP and GEF can both increase GTP hydrolysis but GAP lowers the transition state energy (Scheffzek et al., 1997) while GEF stimulates GDP exchange (Bos et al., 2007). Interestingly, another study found that TP stimulated GTP hydrolysis of atTOC33 only when atTOC33 was in the dimeric form but not in the monomeric form (Oreb et al., 2011).Further analysis found that TP stimulated atTOC33 GTP hydrolysis through interacting with the dimer and increases the nucleotide exchange rate, which suggests the role of TP as a GDP-dissociation inhibitor-displacement factor where each atTOC33 in the dimer acts as GDP-dissociation inhibitor and TP disrupts the dimer (Oreb et al., 2011). The GAP function of TPs is not at odds with a GAD mechanism for TOC34. Based on FRET and PELDOR analysis, Lumme et al (2014) suggest that TOC34 can interact with TPs in a dimeric GDP bound state, and that GTP alters TOC34 conformation to a "looser" structure that would allow for precursor protein import through the TOC75 beta-barrel channel. Thus, TP GAP function may act as a regulatory switch to ensure the specificity of plastid import.

Other discrepancies are found in the GTP hydrolysis measurements of TOC receptors. We compared the crystal structure of TOC34 to the well-studied small GTPase Ras (Bourne et al., 1991; Sun et al., 2002a) (Figure 3-4, 3-5). In the canonical Ras superfamily GTPase mechanism, the carbomyl oxygen of Gln-61 activates a water molecule for nucleophilic attack on the γ-phosphate of GTP(Bourne et al., 1991). Classically, the highly conserved Thr-35 residue of Ras is responsible for interaction with the γ-phosphate of GTP, yet TOC34 family members lack this residue (Bourne et al., 1991; Sun et al., 2002a). Furthermore, In the G3 motif, TOC34 has the sequence motif Asp93-X-X-Gly-Leu97 instead of the canonical Asp57-X-X-Gly-Gln61 of Ras. Interestingly, Ras Gln61 has been established to be essential for GTP hydrolysis activity, so the absence of this residue implies that TOC34 has a modified hydrolysis mechanism (Bourne et al., 1991). Previous work has taken advantage of sequence alignment between the two proteins to predict residues directly involved in the catalytic mechanism of psTOC34 (Sun et al., 2002a). Several residues have been hypothesized to be the catalytic residues of TOC34 (Aronsson et al., 2010). However, mutagenesis in the

proposed psTOC34 “active site” has met with little success: mutants retain WT GTP hydrolysis activity (Aronsson et al., 2010). Furthermore, while one laboratory found that regardless of concentration, monomer or homodimer, psTOC34, atTOC33, atTOC34 and atTOC159 maintain similar GTP hydrolytic rates (Reddick et al., 2007a), another laboratory found that the dimers of psTOC159 and atTOC33 hydrolyze GTP faster than their monomers (Yeh et al., 2007). These differences could be explained by two reasons. First, GTP hydrolysis and dimerization are diminished over time as can be seen in both atTOC33 and psTOC34 (Koenig et al., 2008a; Yeh et al., 2007). Second, the method based on the capture of TOC receptor on a surface may change the monomer-dimer equilibrium compared to what is seen in solution.

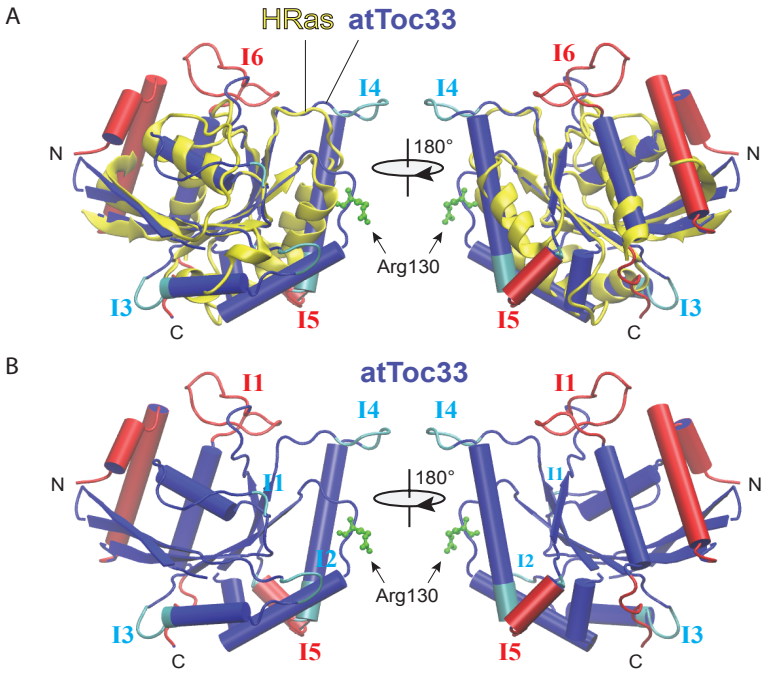
When the Arg130 of atTOC33 and the Arg133 of psTOC34, which were proposed to function similarly to the Arg finger of Ras GAP (Kessler and Schnell, 2002) were mutated to Ala, it was clear that these mutants were unable to dimerize (Koenig et al., 2008a; Lee et al., 2009b; Reddick et al., 2007a; Yeh et al., 2007). While two reports found that atTOC33 mutant (R130A) was able to hydrolyze GTP at 0.6-1 fold of the wild-type level (Koenig et al., 2008a; Lee et al., 2009b), two reports showed that psTOC34 mutant (R133A) hydrolyzed GTP at much lower rate, much less than 0.3 fold of the wild-type level (Koenig et al., 2008a; Reddick et al., 2007a). These results may indicate a structural difference between atTOC33 and psTOC34 functions. While psTOC34 has a high dimerization affinity (Koenig et al., 2008a; Reddick et al., 2007a), atTOC33 has lower affinity for dimerization (Koenig et al., 2008a; Oreb et al., 2011) and is able to support GTP hydrolysis in monomeric form (Koenig et al., 2008a; Lee et al., 2009b).

In summary, current analysis of the TOC GTPases system benefits from a wealth of previous research. We can take advantage of previously established TOC GTP hydrolysis studies and GTPase-inactive mutants to set up a robust system to test our model shown in Figure 3-1. We used a combination of QM/MM modeling, kinetic analysis and crosslinking/FRET interaction studies to build our understanding of the TOC34 GTP hydrolysis mechanism as well as how TOC34 interacts with itself and with the TP. Together, these analyses can be used to gain insight into the mechanism that regulates protein import into the plastid.

Figure 3-4: Structural domains of TOC33/34 with RAS

A. Structural alignment of atToc33 (PDB 3BB4) and human p21Ras (HRas; PDB 5B2X) using jFATCAT tool on PDB server. Paraseptin GTPase atToc33 shows expansion segments (in red and cyan) in comparison to Septin GTPase HRas (in yellow). I1-I5, sequence inserts; G1-G5, GTP/GDP-binding protein motifs.

B. Structure of atToc33 as in panel A but without HRas.


A
HRas
atToc33
I6
I4
I4
I6
N
180°
N
Arg130
I3
I5
C
I5
I3
C
B
atToc33
I1
I4
I4
I1
N
I1
180°
I1
N
I2
I2
Arg130
I3
I5
C
I5
I3
C

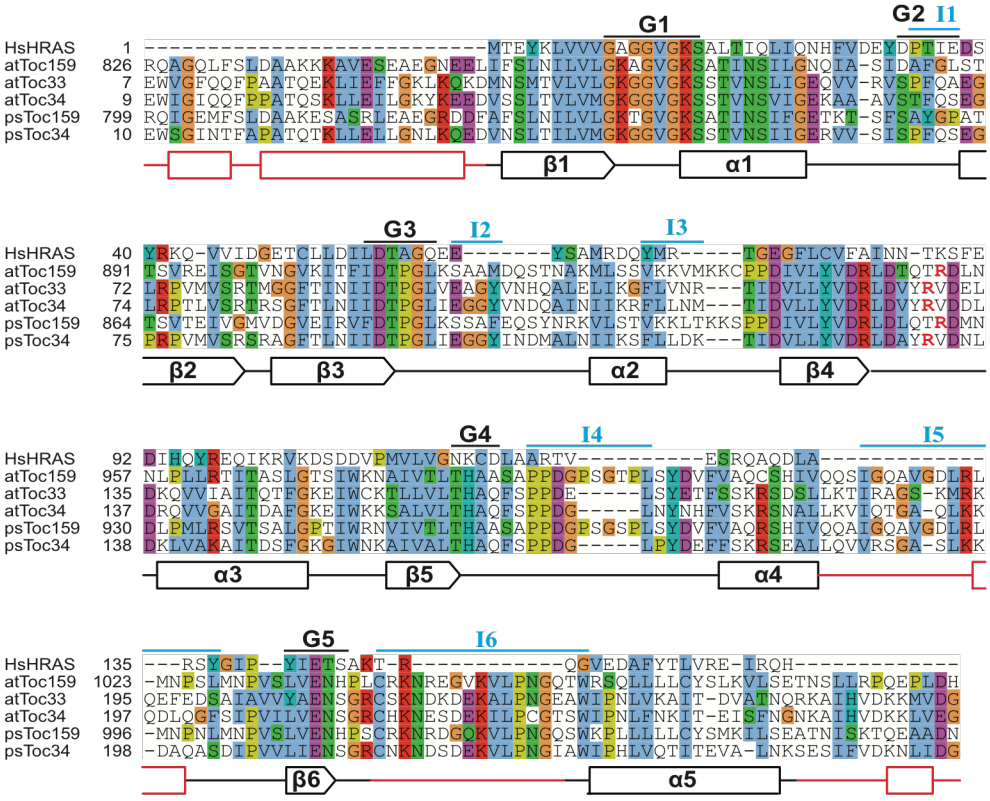


Figure 3-5: Alignment of TOC33/34

Alignment of TOC GTPases with HRas. Top label indicates features of TOC GTPases as in Figure 3-4. The secondary structure depiction at the bottom was derived from atToc33 (PDB 3BB4), psToc34 (PDB 3BB1) and HRas (PDB 5B2X).

3.2 Results

3.2.1 QM/MM and molecular modeling

We collaborated with Dr. Hong Guo's laboratory (University of Tennessee) for our computational experiments. *Ab initio* QM/MM molecular dynamics simulations were performed in CHARMM/GAMESS-US (Figure 3-6). In the current study, we have combined QM/MM calculations and molecular modeling to propose an alternate psTOC34 catalytic site. Based on the QM/MM results, we have employed site directed mutagenesis and highly sensitive GTP hydrolysis assays to characterize the catalytic mechanism of the psTOC34 GTPase. Moleulcar modeling simulations suggest that psTOC34 Glu-73 acts to align a water molecule for nucleophilic attack on the γ-phospate of bound GTP, hydrolyzing the bond and leading to the formation of GDP and free phosphate (Figure 3-6). Suprisingly, this information implies that Glu-73 is a central component of the psTOC34 catalytic mechanism. Previous work had not implicated Glu-73 (Aronsson et al., 2010; Reddick et al., 2007b; Sun et al., 2002a), so we focused our site-directed mutagenesis studies in this region. Importantly, Glu-73 is highly conserved and located within the switch I region of the canonical G2 domain, so Glu-73 appears to be a logical candidate for contribution to the catalysis mechanism (Figure 3-1 and 3-2). We moved our studies into *in vitro* biochemistry to test the hypotheses developed through computational work. We performed site directed mutagenesis to produce a variety of psTOC34 mutants, focusing on both Glu-73 as well as the amino acids structurally close to the proposed active site. We aimed to produce mutants with a variety of effects - from nearly WT GTPase activity to inactive - to carefully probe the hypothesized catalytic mechanism.

3.2.2 Production of psTOC34 mutant proteins

The truncated soluble, GTPase domain of psTOC34 mutant proteins were purified via a recombinant *E.coli* based expression system that is well-established in the Bruce lab (Figure 3-7A) (Reddick et al., 2007b). We used a 6X HIS-tag to isolate TOC34 with metal affinity chromatography. Figure 3-7B lists the array of mutants used in this study. We selected a variety of mutations to alter the physicochemical properties of the active site,

by removing specialized side chains (Ala mutations – E73A, Q71A), adding or removing charge (Q71R, Q71E, E73Q) or adding hydroxylated amino acids (E73S). We also combined mutations to make double and triple mutants (E73G/G74E, Q71E/E73D, Q71E/E73S, Q71E/E73G/G74E). We verified the identity of the mutant proteins using MALDI-TOF mass spectrometry (Figure 3-7B). We tested the global secondary structure using circular dichroism (CD) to ensure that mutagenesis did not alter global protein structure (Figure 3-7C-D). CD spectra were deconvoluted and the protein composition was compared to the WT crystal structure (Figure 3.7D) (Sun et al., 2002a).

3.2.3 Enzymology of psTOC34 mutant proteins

Highly sensitive GTP hydrolysis assays using [γ-^{32}P] labeled GTP were performed on each mutant to test its GTP hydrolytic rate compared to the WT psTOC34 protein, and negative control R133A psTOC34, which is catalytically inactive (Reddick et al., 2007b). We used a protocol first developed in the Bruce lab for these assays; data was collected as Counts per minute (CPM) and converted to enzyme activity units (Reddick et al., 2008b; Reddick et al., 2007b). This data was fit to the Michaelis-Menton equation to calculate V_{max} and K_m kinetic parameters.

We focused our initial studies on both single (Figure 3-8) and double mutants (Figure 3-9). We generated both V_{max} and K_m kinetic parameters for each mutant protein and analyzed their similarity to WT. (Figure 3-10). Importantly, predicted inactive Ala mutation of Glu 73 (E73A) shows near-WT hydrolysis rates, suggesting that the QM/MM prediction is not an accurate depiction of the TOC34 active site. However, unexpectedly, several mutants, particularly E73Q, show increased catalytic rates. It is important to note that the K_m value for WT psTOC34 is slightly higher than previously published results (Figure 3-10) (Reddick et al., 2007b). This is likely due to variations in the experimental procedure as well as variations in purity of the commercially purchased [γ-^{32}P] labeled GTP.

3.2.4 Interaction studies using chemical crosslinking and FRET

We investigated the recognition of a model transit peptide (SS-tp, Rubisco small-subunit transit peptide) with the recombinant form of the cytosolic domain of psTOC34 used in our GTPase analyses (Figure 3-5). We took advantage of the single conserved

cysteine on TOC34 (Cys215) to facilitate homobifunctional crosslinking or attachment of sulfhydryl reactive fluorochromes for FRET. Previously, we created a series of 14 different solvent-exposed TOC34 single Cys mutants, which allow for fluorochromes or crosslinkers to be placed in different locations on the surface of TOC34 (Figure 3-11). Furthermore we generated several cysteine mutations selectively placed along the length of the transit peptide (S3C; S34C; G52C) (Figure 3-11). The collection of both TOC34 and TP Cys mutants provide a range of sulfhydryl-reactive locations so that we can fully probe the interaction site of TOC34-TP.

3.2.4.1 Homobifunctional crosslinking of TOC34

WT TOC34 has a defined, well-studied profile under conditions of homobifunctional crosslinking using BMH and visualized with SDS-PAGE. In the absence of crosslinker, TOC34 is found as a monomeric species, whereas in the presence of BMH, a second species at about 58kDa is observed. The size shift indicates that two TOC34 monomers are linked via their respective Cys215 residue, which is located at the dimeric interface (Sun et al., 2002a). BMH crosslinking covalently joins the monomers, since reducing agent (DTT) in sample loading buffer does not affect their gel mobility. However, TOC34 remains in a monomeric form in the presence of WT TP, even in the presence of crosslinker (Reddick, 2010). This suggests that TP directly interacts with TOC34 and affects the homodimeric interface. Furthermore, negative control R133A mutant TOC34 remains in a monomeric form which indicates that mutation of R133A alters the homodimeric interface. We used WT and R133A as control proteins in our chemical crosslinking studies. For our preliminary analysis, we used BMH crosslinker plus either ATP, GTP or GDP to test the effect of active site mutagenesis on TOC34 homodimerization. Our results are summarized in Figure 3-12. Our crosslinking results somewhat reflect our enzymology analysis; for example, mutants with low enzymatic activity, such as Q71R, E73D and R133A all show reduced homodimerization. However, our BMH analysis did not explain why some mutants show an increase in catalytic rate. CuP is an oxidizing reagant that is a catalytst for oxidation of –SH groups on single Cys residues.

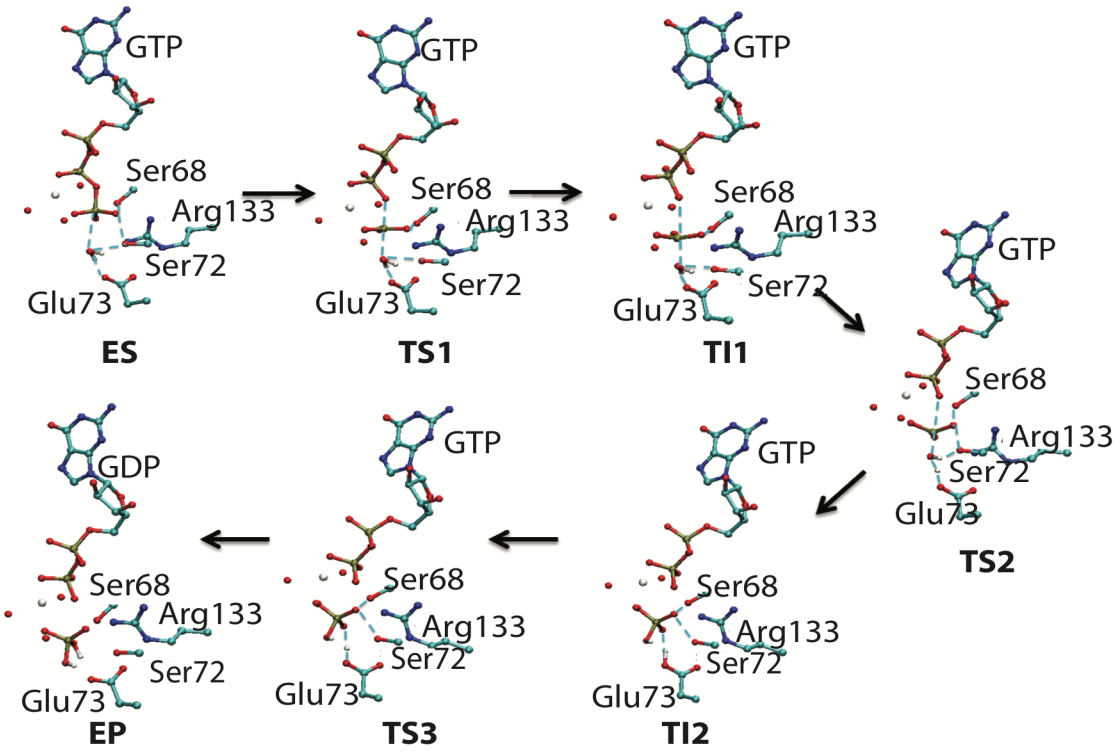
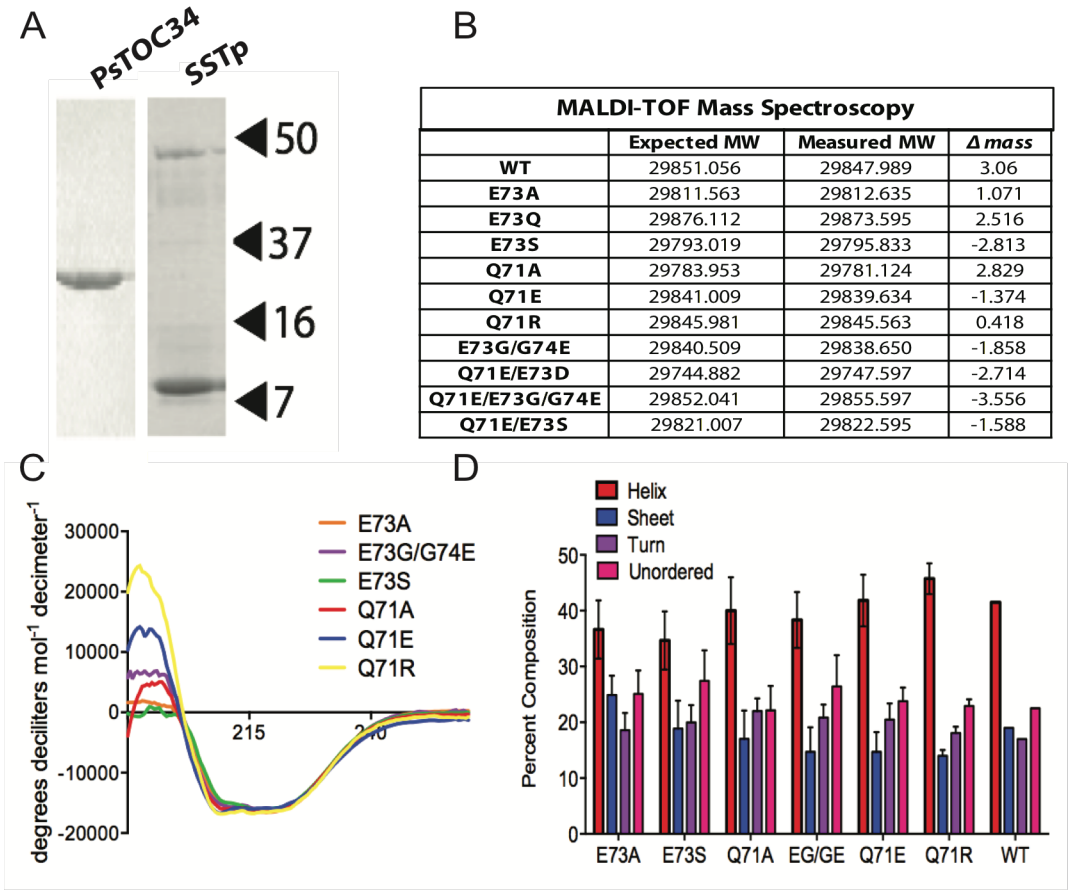


Figure 3-6: QM/MM modeling of TOC34 active site

Proposed GTP hydrolysis mechanism. QM/MM MD simulations were performed by CHARMM/GAMESS-US. Figures were plotted by VMD and residues are shown by ball and stick format. QM/MM MD simulation was performed in collaboration with Dr. Hong Guo's laboratory (University of Tennessee)

Figure 3-7: *E.coli* based recombinant expression and verification of TOC34 proteins

A. psToc34 was expressed in BL-21 RIPL *E. coli* and purified using Ni-NTA resin. Transit peptide was purified using chitin affinity chromatography. The example Toc34 eluate (lane 1) and transit peptide eluate (lane 2) was separated on a Tris-Tricine Gel and visualized by Coomassie Blue staining. B. MALDI-TOF was performed on a Bruker Daltonics MicroflexTM mass spectrometer using Bruker Daltonics Protein Standards II as external and internal standards. Table summarizes the calculated molecular weight and the +1 charge state. Δ mass was calculated using the FindPept tool, www.expasy.org. C. Representative circular dichroism data obtained from psToc34 mutants D. Secondary structure estimation of psToc34 mutants based on CD data deconvoluted with CDPRO. Secondary structure content of WT psToc34 was determined using the STRIDE program with Chain B (PDB ID 1H65).



MALDI-TOF Mass Spectroscopy			
	Expected MW	Measured MW	*Δ mass*
WT	29851.056	29847.989	3.06
E73A	29811.563	29812.635	1.071
E73Q	29876.112	29873.595	2.516
E73S	29793.019	29795.833	-2.813
Q71A	29783.953	29781.124	2.829
Q71E	29841.009	29839.634	-1.374
Q71R	29845.981	29845.563	0.418
E73G/G74E	29840.509	29838.650	-1.858
Q71E/E73D	29744.882	29747.597	-2.714
Q71E/E73G/G74E	29852.041	29855.597	-3.556
Q71E/E73S	29821.007	29822.595	-1.588

Figure 3-7 Continued

Figure 3-8: Enzymology of TOC34 single mutants

γ-32P labeled GTP was used to study GTP hydrolysis of WT psToc34 and psToc34 mutant proteins; the kinetic data for each single mutant is labeled on each graph. Linear regression of CPM vs Time data (decreasing slope is proportional to increasing [GTP]) yields data plotted using the Michaelis-Menton equation, representing both V_{MAX} and K_M. Bottom right panel represents the combined kinetic data for psToc34 single mutants compared to WT psToc34 expressed as % GTP hydrolysis of WT psToc34.

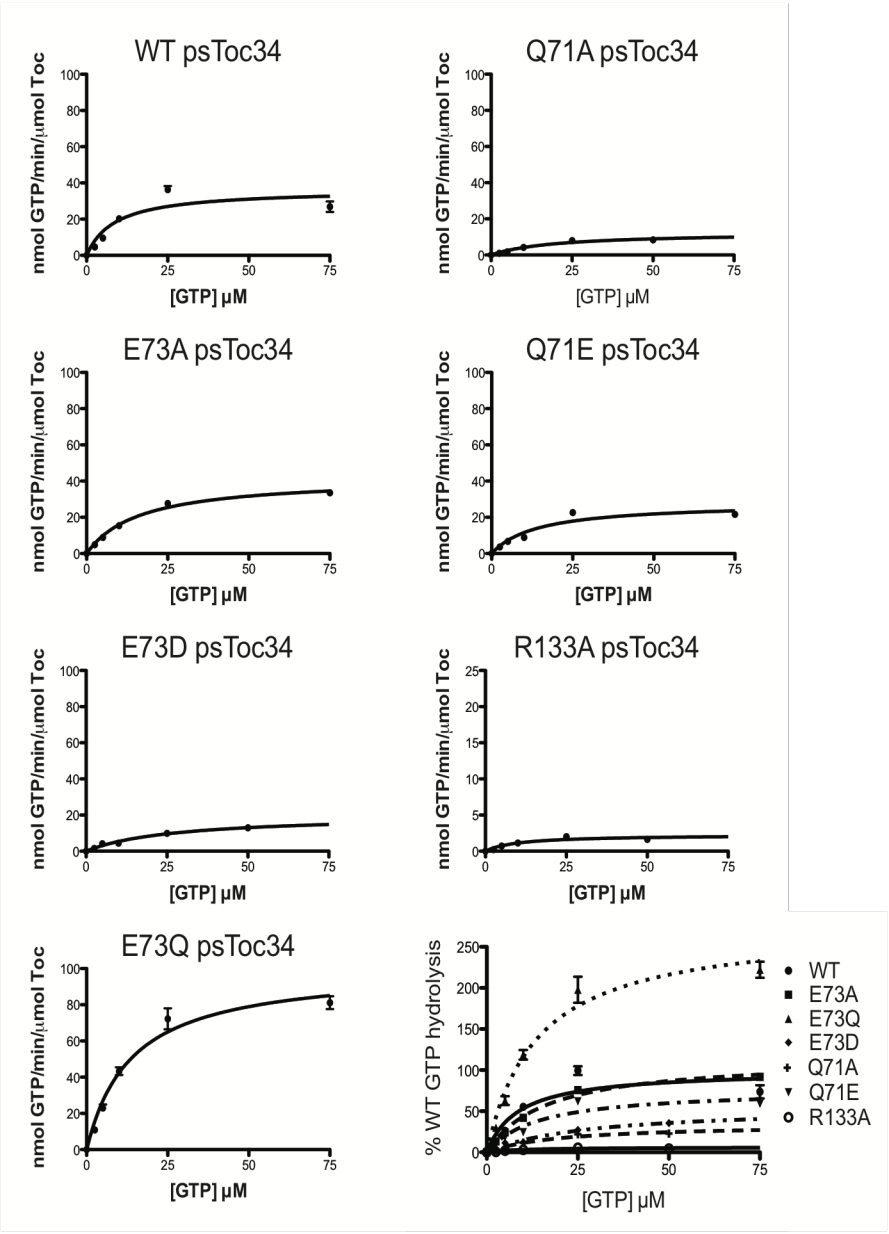


Figure 3-8 Continued

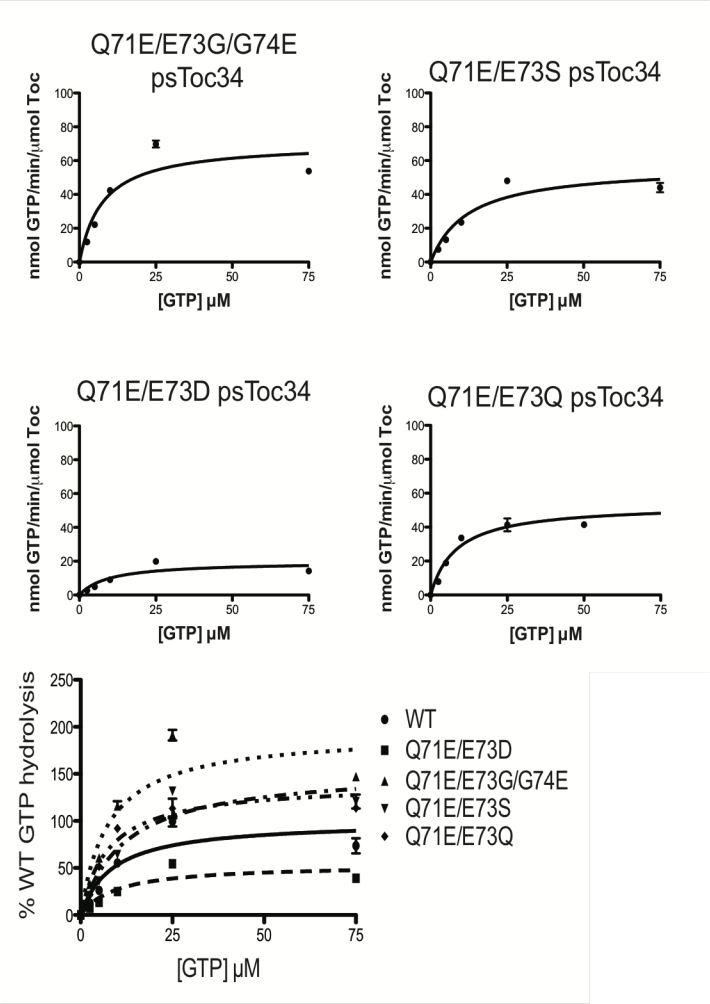


Figure 3-9: Enzymology of TOC34 double and triple mutants

γ-32P labeled GTP was used to study GTP hydrolysis of WT psToc34 and psToc34 mutant proteins; the kinetic data for each mutant is labeled on each graph. Linear regression of CPM vs Time data (decreasing slope is proportional to increasing [GTP]) yields data plotted using the Michaelis-Menton equation, representing both V_{MAX} and K_M. Bottom panel represents the combined kinetic data for psToc34 double mutants compared to WT psToc34 expressed as % GTP hydrolysis of WT psToc34.

TABLE OF ENZYMATIC PARAMETERS			
	V_{MAX} (nm/min/µmol)	K_M (µmol)	% WT GTP hydrolysis
WT psToc34	36.5 ± 5.2	8.8 ± 3.8	100%
E73A psToc34	42.1 ± 1.7	16.4 ± 1.7	115%
E73D psToc34	20.2 ± 1.8	27.6 ± 5.2	55%
E73Q psToc34	100 ± 6.4	13.7 ± 2.4	275%
Q71A psToc34	12.5 ± 1.5	20.1 ± 5.5	34%
Q71E psToc34	28.1 ± 3.2	14.3 ± 4.4	77%
R133A psToc34	2.3 ± 0.3	10.1 ± 3.9	6%
Q71E/E73D psToc34	19.6 ± 3.0	9.7 ± 4.4	54%
Q71E/E73G/G74E psToc34	70.7 ± 8.0	7.6 ± 2.7	194%
Q71E/E73S psToc34	56.8 ± 5.8	12.1 ± 3.5	156%
Q71E/E73Q psToc34	51.3 ± 4.5	7.6 ± 1.9	141%

Figure 3-10: Enzymatic Parameters

Table of measured kinetic parameters for WT psToc34 and single-/double-/triple- active site mutants.

CuP is appealing for our studies because it does not have a spacer arm, thus allowing us to probe very short range interactions. We started this analysis using CuP to crosslink TOC34 and TP. We focused our analysis on WT, increased catalysis mutant E73Q, and reduced catalysis mutants E73D and Q71R (Figure 3-13). We find that both E73D and Q71R show a decreased homodimerization. E73Q shows a somewhat decreased monomerization sensitivity in the presence of TP. Thus, homodimerization may serve as an output to understand changes in TOC34 catalytic rates. However, chemical crosslinking only allows us to take a biochemical snapshot of protein behavior. To develop a more dynamic understanding of TOC34 activity, we developed FRET protocols to assay TP and TOC34 interaction.

3.2.4.2 Development of FRET protocols for TP-TOC34 studies

For FRET experiments, single Cys residues were modified with sulfhydryl-reactive fluorochromes such as ATTO-Tec 495, 590, and 700 (Figure 3-5 and Figure 3-14). To optimize our experiments, we focused on developing protocols to increase fluorochrome labeling efficiency. This is critical for accurate experimental results, because the molar ratio of labeled protein is required to generate distance information (Figure 3-15). We used ATTO-Tec manufacturer's protocols and the degree-of-labeling equation (provided by ATTO-Tec) to assess our protein labeling efficiency. We altered Cys reducing conditions and labeling time to optimize labeling. Seperating labled from unlabeled protein is a technical limitation of this assay; because the fluorochromes are small relative to the size of the protein. Furthermore, optimizing TP labeling is difficult due to the small size of TP. TOC34 can be seperated from unconjugated dye using a size exclusion column, but retention of TP in commerical desalting columns affects protein yield.

We started with preliminary proof-of-principle experiments to test whether the FRET system could be used for TOC34-TP interaction. We generated interaction data using the donor 495 labeled TOC34 (Cys215) protein and acceptor 590-labeled WT (Cys57) SStp and 700-labeled WT SStp. The FRET measurement was sensitive to protein concentration; we could only detect interaction using the 495-590 label combination, suggesting that the interaction is more long-range (Figure 3-16). We

extended our analysis by using S113C mutant TOC34, to move the donor fluorochrome to the opposite side of the TOC34 protein (Figure 3-18). We could not detect 495-590 or 495-700 FRET using this mutant (Figure 3-17B). Using our preliminary data in combination with previous studies, we developed a simple "map" of the possible interaction site(s) for TP-TOC34 (Figure 3-19). Now that we have established the methodology for FRET and chemical crosslinking, we can use an array of Cys mutants to generate a more resolved picture of TP-TOC34 interactions.

3.3 Discussion

We began this project with the aim to develop the mechanistic understanding of the psToc34 GTP hydrolysis cycle. Using QM/MM and molecular modeling, we hypothesized that E73A would have a disruptive effect on TOC34 GTPase activity, yet our experimental data shows that E73A has a near WT kinetic rate. Although this does not rule out a possible role for E73 in the catalytic mechanism, it suggests that TOC34 does not have the same mechanism predicted by the QM/MM calculations.

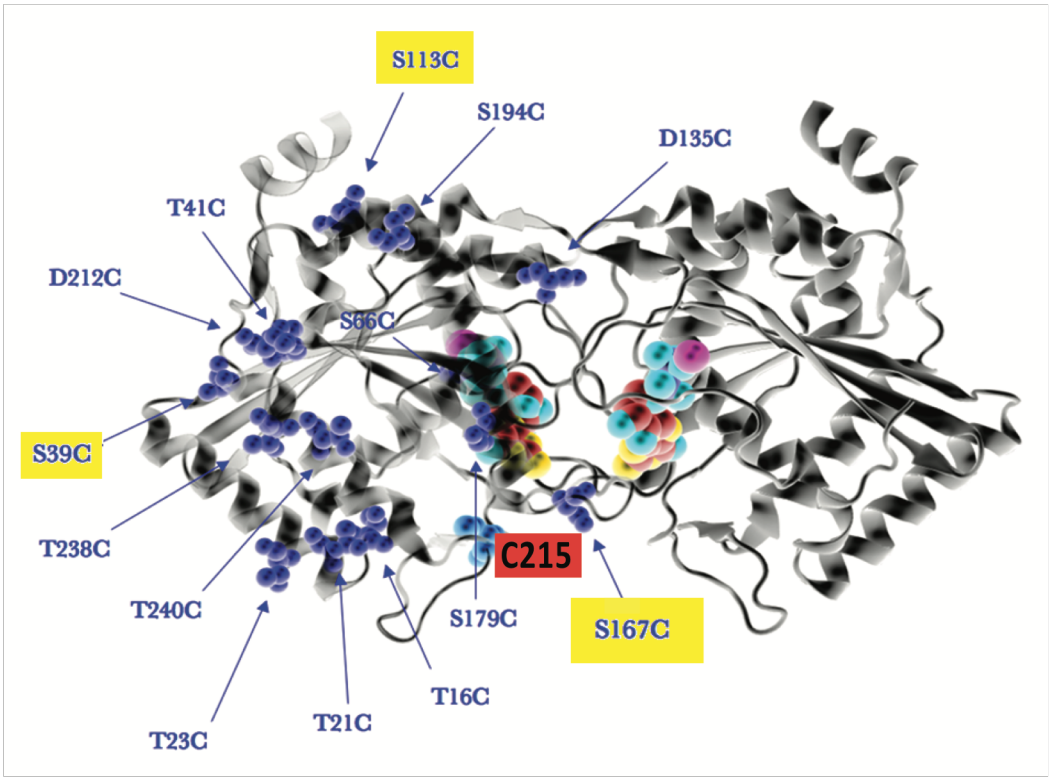
Unexpectedly, we have found several mutants that show increased catalytic ability; however, the mechanism by which these mutations affect psTOC34 activity remains unclear. To gain mechanistic insight, we have aligned psTOC34 with the canonical human GTPase, p21Ras (Figure 3-18). The activitly of small GTPases like Ras are modulated by interacting partners known as Guanine Exchange Factors (GEFs, which turn "on" the GTPase), and GTPase Activating Proteins (GAPs) which turn "off" the GTPase (Bourne et al., 1991). Previous data has established that WT TP has a GTPase activating function for psTOC34 and increases its catalytic rate by 2-fold and may have a GAP-like role in TOC34 function (Reddick et al., 2007b). However, the mechanism of transit peptide GAP activity is unknown. We speculate that mutations that increase psTOC34 GTPase rate may be mimicking the structural changes of TP interaction with TOC34. We plan to test this hypothesis by performing GTPase assays with mutant proteins in the presence of TP, and we expect that these "GAP locked" mutants will be insenstive to TP stimulation (Figure 3-20). By using a combination of kinetics and molecular modeling, we expect to gain further insight into the TOC34 GTP hydrolysis mechanism.

Figure 3-11: psToc34 and RuBisCO SStp Cysteine Mutants

A. psToc34 crystal structure (PDB 1H65) dimer rendered in PyMol. GDP is shown as a space filling model. Solvent exposed cysteine mutations highlighted in blue space-filling representation. These cysteine mutations serve as attachment points for fluorochromes to be used for FRET experiments. Mutants S39C, S167C, and S113C will be used as possible negative controls. WT C215 is highlighted in red.

B. Protein sequences of wild type, C57Y, and the double mutants, with mutations underlined along length of sequence. C. A cartoon schematic of the RuBisCO SStp with Cys mutations shown along the length of the TP. These Cys residues will be used to attach fluorochromes to investigate different regions of the transit peptide and their interactions with psToc34.

A



B

Wild Type:
MASSVLSSAA-VATRSNVAQA-NMVAPFTGLK-SAASFPVSRK-QNLDITSIAS-NGGRVQCMQP-G
1.....11.....21.....31.....41.....51.....61
C57Y: (No-Cysteine Mutant)
MASSVLSSAA-VATRSNVAQA-NMVAPFTGLK-SAASFPVSRK-QNLDITSIAS-NGGRVQ**Y**MQP-G
1.....11.....21.....31.....41.....51.....61
Double Mutants: (Single-Cysteine Mutants)
MA**C**SVLS**C**AA-VAT**CC**N**C**AQA-**C**MVAPF**C**GLK-**C**AA**C**FPV**CC**K-QNLD**CC**SIAS-N**C**GRVQ**Y**MQP-G
1.....11.....21.....31.....41.....51.....61

C

Figure 3-11 Continued

Figure 3-12: BMH based crosslinking of TOC34 homodimers
Gel images show psTOC34 WT and mutant proteins crosslinked using homobifunctional, sulhydryl reactive bismaleimidohexane and separated using SDS-PAGE gel electrophoresis. were crosslinked using Quantification bars represent crosslinking visualized using SDS-PAGE. Crosslinking was measured via densitometry analysis (QuantOne software) and normalized to control WT amounts. Bar graph depicts the Dimer:Monomer ratio (WT normalized to 100%).

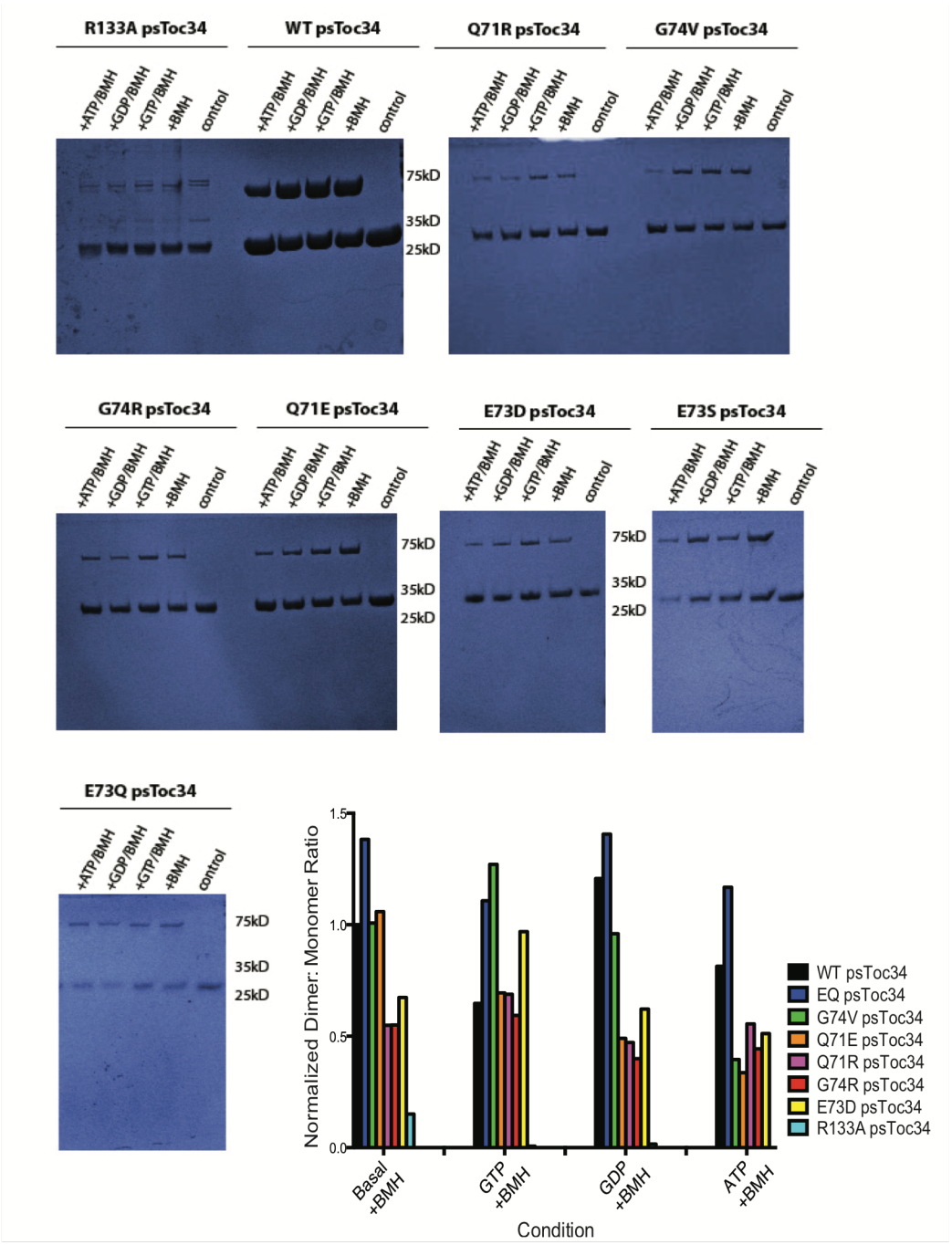


Figure 3-12 Continued

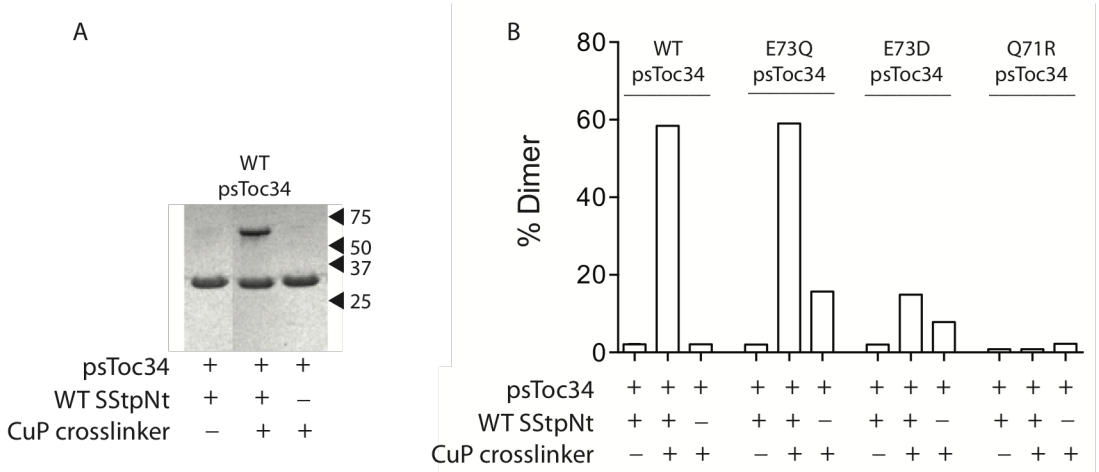


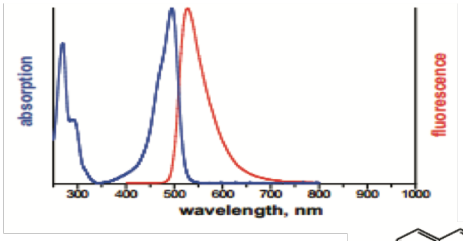
Figure 3-13: CuP based crosslinking of TOC34 homodimers

psToc34 WT and mutant proteins were crosslinked using copper-phenanthroline crosslinker (CuP). A. Representative image of CuP mediated crosslinking of WT Toc34 separated using SDS-PAGE. Conditions are denoted by the symbols (+) and (-). B. Quantification of CuP crosslinking of WT and mutant psToc34 active site mutants, before and after incubation with SStp as denoted by symbols (+) and (-).

Figure 3-14: Fluorochrome Properties and Structure

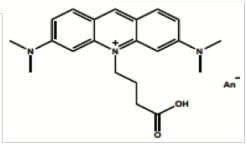
A**.** Structure and absorption and emission spectrums of the chosen fluorochrome: Atto-tec 495. Table shows optical data of the carboxy derivative (in water) for this fluorochrome. B**.** Structure and absorption and emission spectrums of the chosen fluorochrome: Atto-tec 590. Table shows optical data of the carboxy derivative (in water) for this fluorochrome. C**.** Absorption and emission spectrums of the chosen fluorochrome: Atto-tec 700. Table shows optical data of the carboxy derivative (in water) for this fluorochrome. The structure for Atto-tec 700 is undisclosed.

A

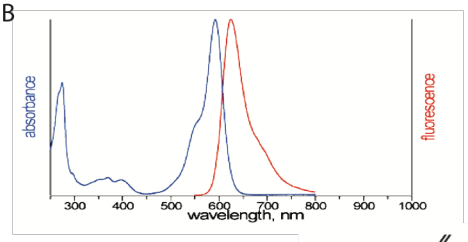


λ_{abs}	=	495 nm
ε_{max}	=	8.0 x 10^4 M^{-1} cm^{-1}
λ_{fl}	=	527 nm
η_{fl}	=	20 %
τ_{fl}	=	1.0 ns
CF_{260}	=	0.57
CF_{280}	=	0.39

Atto-tec 495

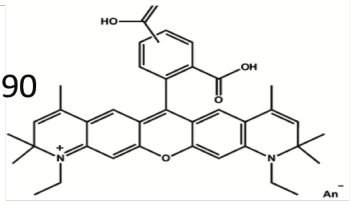


B

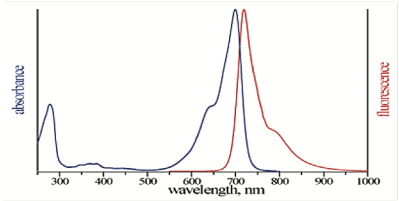


λ_{abs}	=	594 nm
ε_{max}	=	1.2 x 10^5 M^{-1} cm^{-1}
λ_{fl}	=	624 nm
η_{fl}	=	80 %
τ_{fl}	=	3.7 ns
CF_{260}	=	0.42
CF_{280}	=	0.44

Atto-tec 590



C



λ_{abs}	=	700 nm
ε_{max}	=	1.2 x 10^5 M^{-1} cm^{-1}
λ_{fl}	=	719 nm
η_{fl}	=	25 %
τ_{fl}	=	1.6 ns
CF_{260}	=	0.26
CF_{280}	=	0.41

Atto-tec 700

Structure of Atto-tec 700 is not disclosed

Figure 3-14 Continued

Figure 3-15: Optimization of Toc34-Atto495 Labeling Efficiency

A. The degree of labeling (DOL, dye-to-protein ratio) obtained by labeling procedures (described by Atto-Tec) can be determined by absorption spectroscopy. Using the UV-Vis spectrum, the absorbance (Amax) at the absorption maximum (λabs) of the dye and the absorbance (A280) at 280 nm (absorption maximum of protein) are determined. The concentration of the conjugated dye is given by: c(dye) = Amax / εmax • d, where εmax is the extinction coefficient of the dye at the absorption maximum. The protein concentration is determined from its absorbance maximum at 280nm (εprot is the extinction coefficient of the protein at 280 nm). Because all dyes exhibit some absorption at 280 nm, the measured absorbance A280 must be corrected for the contribution of the dye. This is given by Amax • CF280. Calculated DOL can vary by about 20% according to Atto-tec manufacturer's instructions. B. The degree of labeling determined for WT-Toc34 incubated with Atto-495 under various conditions for optimization. TCEP reducing agent was used as recommended by Atto-Tec. Under all conditions, after reducing treatment, label was incubated with protein for 2 hours at room temperature in the dark before separation of unconjugated label using Zeba (Pierce) desalting columns). UV-VIS absorbance spectra were collected to determine the sample absorbance at 280nm and 495nm.

A

$$DOL = \frac{A_{max} / \varepsilon_{max}}{A_{prot} / \varepsilon_{prot}} = \frac{A_{max} \cdot \varepsilon_{prot}}{(A_{280} - A_{max} \cdot CF_{280}) \cdot \varepsilon_{max}}$$

B

Condition	DOL (Atto-495)
WT Toc34 control (no reduction) + Atto-495 (DMSO)	11%
WT Toc34 treated with immobilized TCEP reducing gel + Atto-495 (DMSO)	9%
WT Toc34 treated with immobilized TCEP reducing gel + Atto-495 (DMF)	13%
WT Toc34 treated with 25X molar excess TCEP/desalting + Atto-495 (DMSO)	100%
WT Toc34 treated with 25X molar excess TCEP/desalting + Atto-495 (DMF)	100%

Figure 3-15 Continued

Figure 3-16: Preliminary FRET based interaction studies with TOC34 and TP

A**.** Measured interaction between wild type TOC34 labeled with Atto-tec495 and S3C TP labeled with Atto-tec 590 and Atto-tec 700. B. FRET emissions of 495-labeled wild type Toc34 with 590-labeled wild type SS-tp organized by molar ratio. C. The FRET efficiency of 495-labeled wild type Toc34 and 590-labeled wild type SS-tp and 700-labeled wild type SS-tp vs. the molar ratio of transit peptide to Toc34.

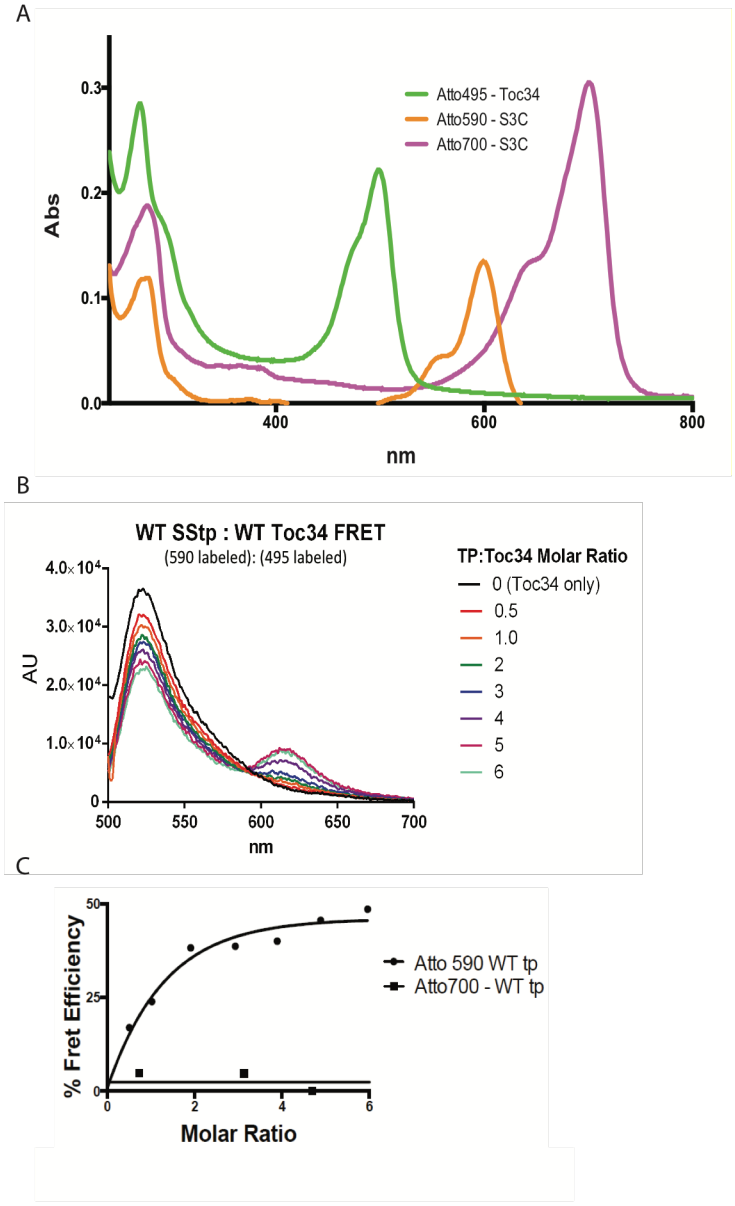
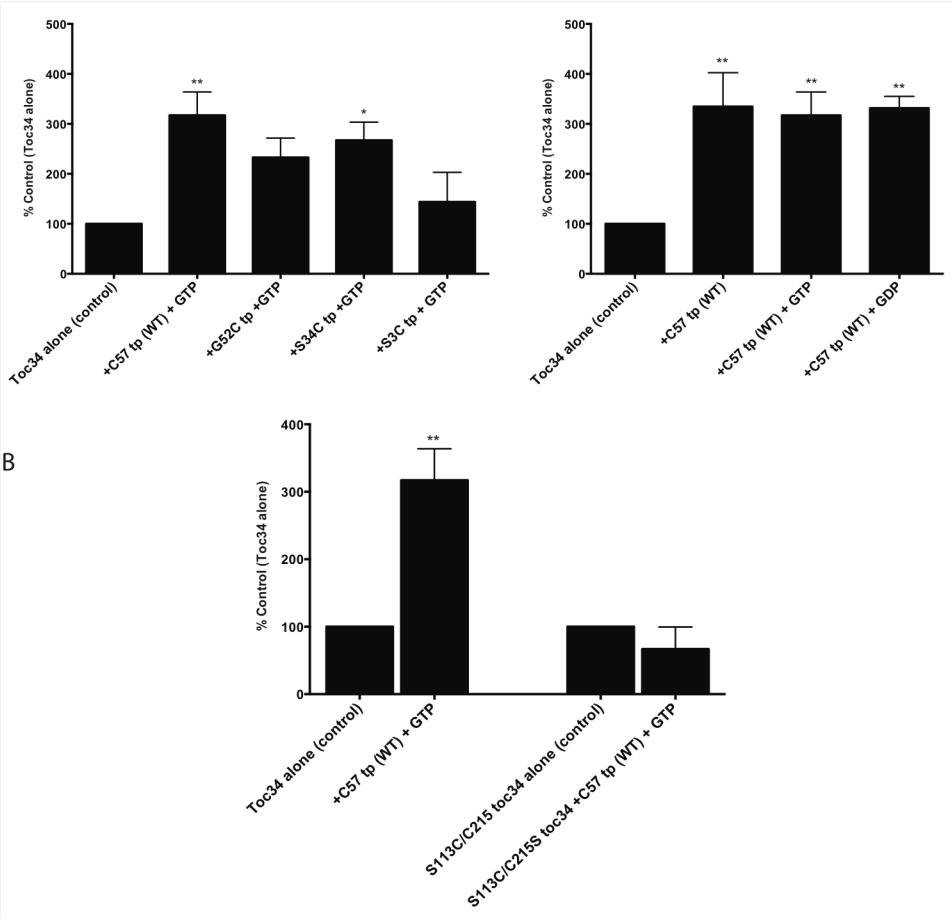


Figure 3-16 Continued

Figure 3-17: FRET based interaction studies of Toc34

Bars show the mean ± SEM FRET signal of WT toc34-Atto 495 incubated with various SS-tp-Atto 590 constructs. Raw fluorescence counts were baseline subtracted and normalized to WT toc34 alone (100%). * and ** indicate conditions where FRET signal was significantly higher than the control value ($p<0.05$ and $p<0.0015$, respectively) One Way ANOVA with Bonferroni Post Hoc Analysis. A. Left panel bar graph shows interaction between WT labeled TOC34 and TP labeled at the N-terminus (S3C), middle domain (S34C), or C-terminus (G52C, WT). Right panel bar graph shows interaction between WT labeled TOC34 and WT labeled TP with basal, GTP or GDP nucleotide added. B. Bar graph show FRET signal between WT labeled TOC34 or S113C/C215S labeled TOC34 and WT labeled TP in the presence of GTP.

A



B

Figure 3-17 Continued

We have performed preliminary FRET experiments with wild type psTOC34 and S113C/C215S psTOC34 as donors and wild type SStpNT, S3C/C57Y, S34C/C57Y, and G52C/C57Y SStpNT. We hope to investigate which region of the TP is bound to the psTOC34 by looking at the anisotropy of the fluorescently labeled transit peptide as well as possible spectral shifts of the fluorochrome. In addition, we have also synthesized 14 different solvent-exposed single cysteine TOC34 mutants. We plan to use these mutants to investigate the region of TOC34 that may interact with SS-tp using Fluorescence Resonance Energy Transfer (FRET). We expect that this project can be expanded to further investigate the dynamics of the TOC complex interactions. A plate reader is being optimized for our FRET experiments and will be used for future experiments due to its ability to support multiple conditions using smaller quantities and substantially less time with comparable results to the fluorometer. Emissions from the acceptor fluorochrome (with auto-activation of the acceptor fluorochrome subtracted) indicate interaction between donor and acceptor. This information will be gathered to determine which regions of transit peptide interact with which regions of TOC34 based on the FRET efficiency of our donor and acceptor pairs. Our proposed model (Figure 3-19) of TP and TOC34 interaction is based on FRET and data using an alternate fluorochrome called MIANS (Reddick, 2010). Taken together, our work suggests that the C-terminus of transit peptide is more important for binding TOC34 than the N-terminus. The binding region shown here of TOC34 is based on similarities with the major histocomatibility complex. Atto-tec 590 labeled WT SS-tp emits due to energy transfer from Atto-tec 495 labeled WT psTOC34. Atto-tec 590 labeled S3C/C57Y SS-tp does not emit, as it is too far from the donor to be excited. The results so far support our proposed model (Figure 3-19) taking into consideration labeling efficiencies, which began low, but are improving greatly with each experiment. We are currently collecting data on additional SStp titration experiments. Beyond these FRET experiments, the cysteine mutants of transit peptide can also be fluorescently labeled and used for anisotropy experiments. In summary, the groundwork of this project can be expanded to dissect the details of the psTOC34 catalytic mechanism, and eventually apply our detailed findings to better understand the import cycle of the TOC translocon.

Figure 3-18: The Active Site of Toc34

Superposition of WT psToc34 (green) and human p21 Ras (blue) crystal structures (PDB 1H65 and 4Q21, respectively). Inset shows a close-up of the alignment of the known Ras catalytic residue (Q61), and residues involved in GTP binding (T35) and γ-phosphate stabilization (G60). psToc34 residues of interest shown in green, and GDP is shown in orange.

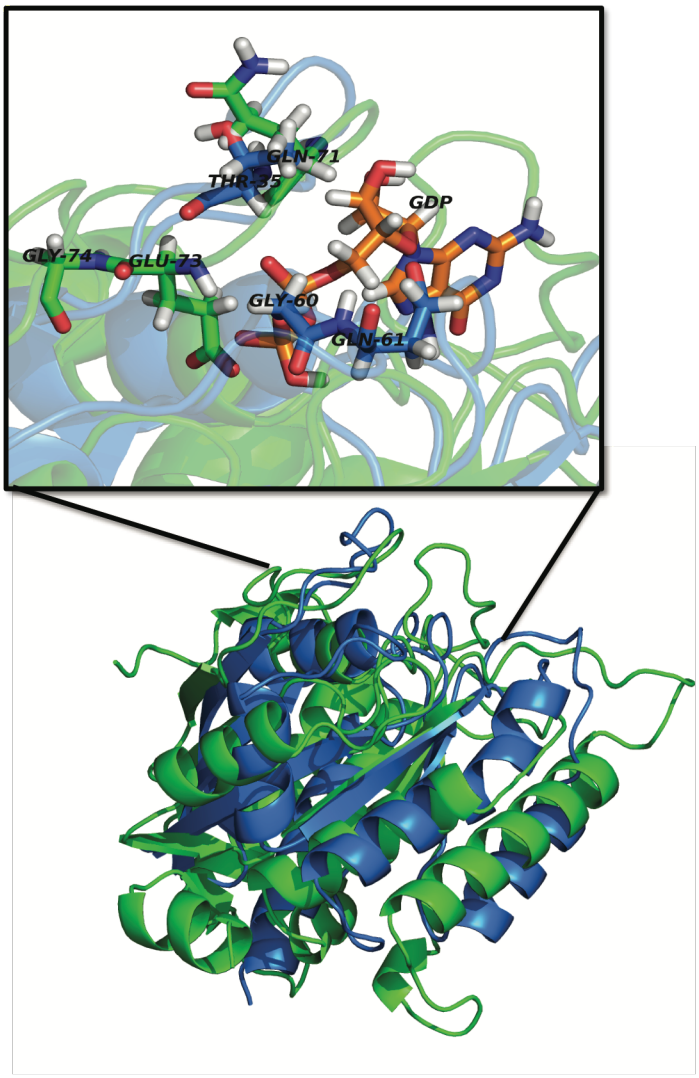


Figure 3-18 Continued

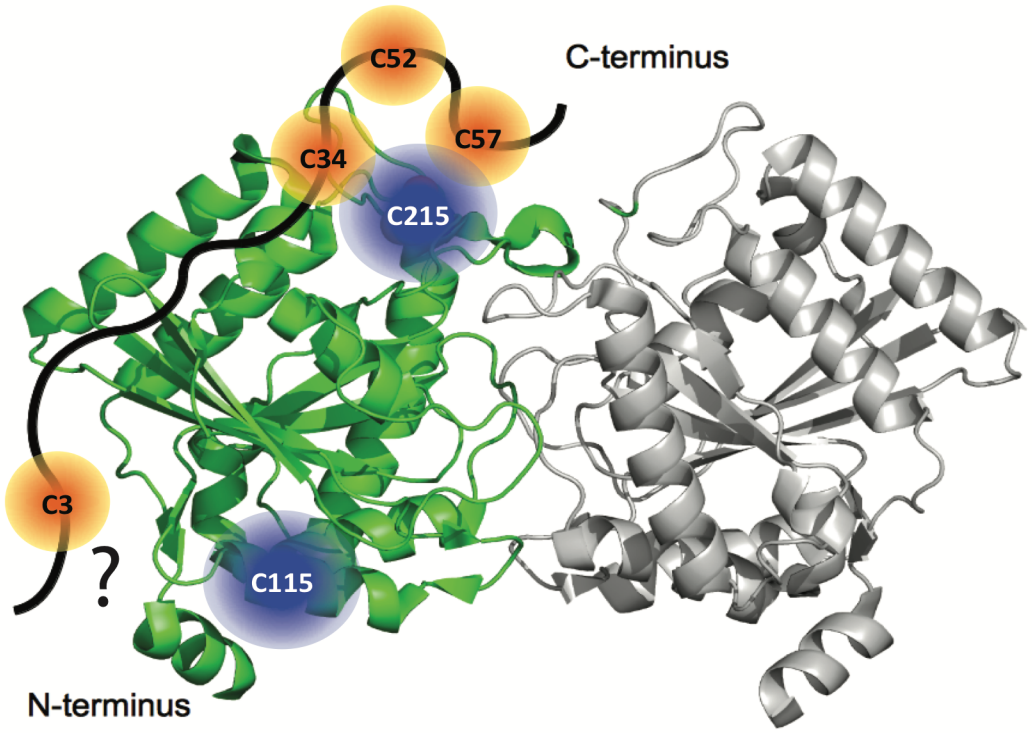


Figure 3-19: Model for Toc interaction with TP

A model of transit peptide and Toc34 interaction based on FRET and MIANS data (Reddick 2010) suggesting the C-terminus of transit peptide is more important for binding Toc34 than the N-terminus. The binding region shown here of Toc34 is based on similarities with the major histocomatibility complex. Atto-tec 590 labeled WT SS-tp emits due to energy transfer from Atto-tec 495 labeled WT psToc34. Atto-tec 590 labeled S3C/C57Y SS-tp does not emit, as it is too far from the donor to be excited.

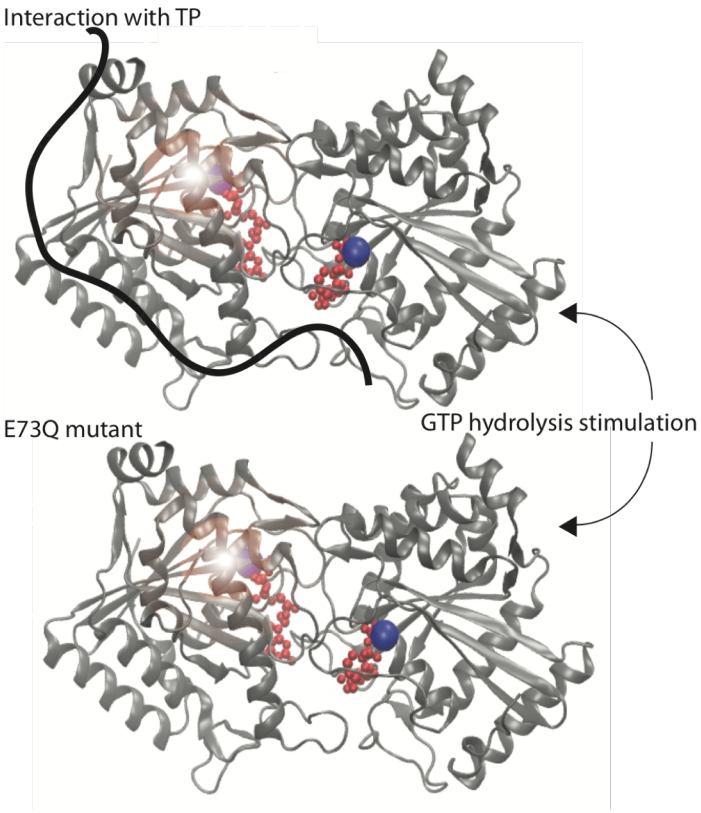


Figure 3-20: Model of E73Q GTP hydrolysis stimulation

Cartoon schematic depicts the proposed stimulation of GTP hydrolysis by either WT TP interaction or E73Q mutation. Additional work will test whether E73Q is “locked” in a conformation similar to productive interaction with WT TP.

CHAPTER 4: FUNCTIONAL ANALYSIS OF SEMI-CONSERVED TRANSIT PEPTIDE MOTIFS

Disclosure:

The work reported in this chapter has been published in a research paper by Holbrook et al (2016). The results primarily generated by other authors are cited in the Figure Legends as appropriate and are included in this chapter only to clarify the findings.

Abstract:

Over 95% of plastid proteins are nuclear-encoded as precursors containing an N-terminal extension known as the transit peptide (TP). Although highly variable, TPs direct the precursors through a conserved, posttranslational mechanism involving translocons in the outer (Toc) and inner envelope (Tic). The organelle import specificity is mediated by one or more components of the Toc complex. However, the high TP diversity creates a paradox on how the sequences can be specifically recognized. An emerging model of TP design is that they contain multiple loosely conserved motifs that are recognized at different steps in the targeting and transport process. Bioinformatics has established that many TPs contain semi-conserved physicochemical motifs, termed FGLK. In order to characterize FGLK motifs in TP recognition and import, we have analyzed two well-studied TPs from the precursor of RuBisCO small subunit (SStp) and ferredoxin (Fdtp). Both SStp and Fdtp contain two FGLK motifs. Analysis of large set mutations (~85) in these two motifs using *in vitro, in organello,* and *in vivo* approaches support a model in which the FGLK domains mediate interaction with Toc34 and possibly other TOC components. *In vivo* import analysis suggests that multiple FGLK motifs are functionally redundant. Furthermore, we discuss how FGLK motifs are required for efficient precursor protein import and how these elements may permit a convergent function of this highly variable class of targeting sequences.

4.1 Introduction

Eukaryotic organisms are defined by their compartmentalization and various organelles. The membranes that define these organelles present a barrier to the selective translocation of proteins from the cytosol into their functional location in the organelle. In plant cells, the plastid represents one of the most complex systems of protein sorting, requiring several translocons located in the three membranes found in this organelle (Andrès et al., 2010; Bruce, 2000b; Jarvis, 2008b; Li et al., 2007; Schleiff et al., 2003b; Schnell et al., 1990). The vast majority of plastid-localized proteins are nuclear-encoded and must be post-translationally imported from the cytosol. These proteins are encoded as precursors with a N-terminal sequence of approximately 55 amino acids called a transit peptide (TP), which is required for efficient recognition and import of target proteins into the chloroplast (Andrès et al., 2010; Bruce, 2000b; Bruce, 2001b; Chotewutmontri et al., 2012; Dobberstein, 1977; Li et al., 2007; Schmidt GW, 1979).

Although it has been well established that the TP is required for recognition and import of precursor proteins by the Translocon of Outer Chloroplast membrane (TOC) and Translocon of the Inner Chloroplast membrane (TIC), neither the mechanistic details of this interaction nor how this interaction promotes an organelle specific recognition are well understood. The mechanism by which proteins are targeted to the chloroplast has been largely determined by careful analysis of *in vitro* import assays using purified chloroplasts (Keegstra, 1989; Keegstra, 1999).Various models have shown that translocation can happen coordinately using ATP hydrolysis or in distinct steps across each envelope membrane (Liu et al., 2014; Scott and Theg, 1996). It is widely hypothesized that the primary targeting step providing organelle specificity is via interactions between the TP and the TOC receptors, TOC34 and TOC159 (Chotewutmontri et al., 2012; Jelic et al., 2002a; Lee, 2009; Reddick et al., 2007b; Sveshnikova et al., 2000). Although its clear that TPs are highly divergent in sequence and length, a universal model of chloroplast protein import will need to account for a common mechanism(s) allowing highly divergent TPs to be recognized (Bruce, 2000b). Previous work from multiple groups has focused on a combination of mutagenesis (Lee, 2002; Pilon et al., 1995b), deletion (Kindle, 1998a, b; Rensink, 1998, 2000) and domain swapping experiments (de Castro Silva Filho, 1996) to identify TP functional domains

(von Heijne et al., 1989b); however, based on TP sequence variation, these results are difficult to extend to TPs in general. Recent work has focused on physicochemical motifs, rather than identification of primary sequence similarity, to identify common modes of TP recognition. We have identified a specific N-terminal TP domain that is required to facilitate interaction with plastid-localized molecular chaperones (Chotewutmontri et al., 2012; Ivey et al., 2000b). This motif is followed by a second physicochemical element that interacts with the TOC receptor proteins and appears to promote TP chloroplast binding, but cannot support import alone. This region is termed 'FGLK,' which is loosely defined as a (1) aromatic amino acid, (2) a turn-inducing or helix breaking amino acid, (3) a small nonpolar amino acid, (4) a basic amino acid, and (5) the absence of negatively charged amino acids Asp or Glu (Chotewutmontri et al., 2012; Lee, 2006, 2009; Pilon et al., 1995b). Thus, previous work suggests that TPs do not share any consensus sequences, but contain specific physicochemical elements that allow them to be recognized by a common import mechanism utilizing a discrete set of Toc GTPases, a single Toc75 complex, and a universal requirement for stromal ATP. The properties of TP physicochemical blocks are likely to be context sensitive, and may behave differently as a function of pH, membrane-like conditions, or the presence of translocon components (Bruce, 1998b; Chotewutmontri et al., 2012; Lancelin et al., 1994a; von Heijne, 1991; Wienk, 1999).

To dissect the contributions of the FGLK domains individually and each of the component amino acids, we have developed a systematic approach using two well-studied TPs. We have generated a comprehensive set of deletion and substitution mutants to evaluate physicochemical elements such as hydrophobicity, flexibility and charge within the FGLK domain. Our approach includes *in vitro* analysis of the interaction of mutant peptides with the isolated cytosolic GTPase domain of TOC34, and *in organello* assays to measure binding and import of mutant peptides with isolated chloroplasts. We have also developed *in vivo* analyses in onion epidermal cells to assess the ability of the cell to target and import mutant peptides into the plastid under physiological conditions. Overall, our results highlight the ability of a set of loosely conserved physicochemical properties of TPs to maintain both the specificity and promiscuity of the plastid import mechanism.

4.2 Results

4.2.1 Alignment of SStp reveals semi-conserved domains

Previous research has suggested that semi-conserved physicochemical motifs have an important functional role in the recognition and import of chloroplast targeted proteins (Chotewutmontri et al., 2012; Ivey et al., 2000b; Lee, 2009; Pilon et al., 1995b). To determine the extent of the FGLK domain presence and conservation, we aligned 328 full-length protein sequences of the small subunit of RuBisCO from 120 green plants (data not shown). Representative sequence alignment of SStp from green algae, moss, rice, *Arabidopsis*, pea and tobacco are shown in Figure 4-1A. Identified consensus sequences from different lineages are highlighted as a logo plot in Figure 4-1B; the height of the black bar denotes the abundance of the consensus residue, and the height of the letter indicates the abundance of that amino acid in the consensus motif. Our analysis identified two FGLK domains, Domain I (DI) and Domain II (DII), which are highlighted in gray (Figure 4-1A and 4-1B). FGLK domains have been implicated in TP import (Chotewutmontri et al., 2012; Lee, 2009; Pilon et al., 1995b), so we focused our studies on the detailed analyses of this physicochemical element.

4.2.2 The semi-conserved FGLK domain is required for chloroplast binding and import

To detect the binding of SStp to the chloroplast surface, we used a combination of flow cytometry, fluorescent microscopy and Far Western Blotting (Subramanian et al., 2001). We used a dual epitope tagged form of SStp denoted H-S-SStp (His_6– S-peptide – SStp, expressed recombinantly in *E. coli* and purified (data not shown)) to directly monitor the effect of the TP without contributions from the mature domain. Although this construct is able to compete with WT TP for preprotein binding, it cannot be translocated into the stroma (Subramanian et al., 2001). Transit peptides with the dual epitope present still function in chloroplast binding and import completion assays with the same efficiency as unmodified peptides (Subramanian et al., 2001). The three C-terminal deletion mutants, H-S-$SStp_{\Delta h}$, H-S-$SStp_{\Delta H}$, and H-S-$SStp_{\Delta a}$, have two, one, or zero FGLK motifs remaining in their sequences respectively (Figure 4-2A).

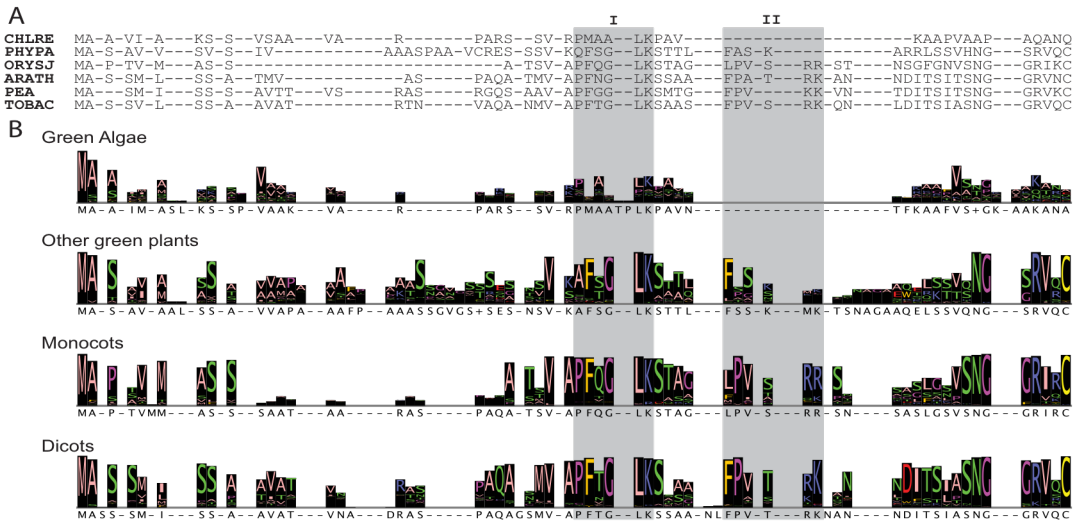


Figure 4-1: Alignment of the small subunit of RuBisCO transit peptides

Two semi-conserved FGLK domains were identified and highlighted among these sequences. A. Representative sequences from different lineage groups. CHLRE, green alga *Chlamydomonas reinhardtii*; PHYPA, moss *Physcomitrella patens*; ORYSJ, rice *Oryza sativa* subsp. *japonica*; ARATH, *Arabidopsis thaliana*; PEA, pea *Pisum sativum*; TOBAC, tobacco *Nicotiana tabacum*. B. Consensus sequences and sequence logos of the alignment separated into lineages. The height of the black bar shows the relative abundance of the consensus residue. The height of amino acid letter in a bar shows the relative abundance of the residue in that column. Bioinformatic analysis performed by Prakitchai Chotewutmontri.

We used a previously established flow cytometric analysis (Subramanian et al., 2001) to measure the amount of H-S-SStp bound to chloroplast. This assay allows direct fluorescent measurement of the amount of H-S-SStp bound to individual chloroplasts via by detection of S-tag of H-S-SStp and using FITC-S-protein (Kim and Raines, 1993). One advantage of FACS is that each organelle is individually measured, instead of measuring an average signal that results from a population of organelles. Thus, this assay permits high-throughput detection, allowing a large number of organelles (>10,000) to be analyzed for statistically significant results. In our previous study, we established controls to specifically identify intact chloroplasts and controls for chlorophyll autofluorescence (Subramanian et al., 2001). Although all 3 mutants have the H-S motif at the N-terminus, only H-S-SStp$_{\Delta 5}$ binds to chloroplasts similar to the WT construct; the other mutants, H-S-SStp$_{\Delta 25}$ and H-S-SStp$_{\Delta 36}$ are negative for FITC staining (Figure 4-2B). Comparison of the H-S-SStp$_{\Delta 5}$ and the H-S-SStp$_{\Delta 25}$, would suggest a key determinant resides between residue G35 and R54; this region contains all of Domain II.

We tested the spatial distribution of the bound H-S-SStp on the chloroplast surface using laser scanning confocal microscopy (LCSM) and an S-protein FITC conjugate to label intact pea chloroplasts after a binding reaction (30 μM ATP) with H-S-SStp constructs (Figure 4-2C). Chloroplasts are depicted in red, and green labeling indicates regions of H-S-SStp interaction on the chloroplast surface. We find that labeling of both H-S-SStp$_{WT}$ and H-S-SStp$_{\Delta 5}$ reveal punctate, patch-like labeling. However, H-S-SStp$_{\Delta 25}$ and H-S-SStp$_{\Delta 36}$ show no labeling, supporting our FACS analysis.

Chloroplast interaction was further verified using far-Western blotting of total chloroplast proteins using the S-protein AP conjugates following binding and re-isolation over Percoll gradient (Figure 4-2D). Both WT and H-S-SStp$_{\Delta 5}$ were easily detected, and as expected H-S-SStp$_{\Delta 25}$ and H-S-SStp$_{\Delta 36}$ do not show chemiluminescent signal (Figure 4-2B-C). Importantly, these results suggest that removal TP length and/or removal of the Domain II motif prevents productive interaction with the chloroplast outer membrane.

Figure 4-2: Truncation of the transit peptide C-terminus abrogates import and chloroplast interaction

A. Schematic showing the sequence and domains of WT His-S-SStp and deletion mutants B. Graphical analysis of FACS study of chloroplast binding of H-S-SStp constructs tagged with S-protein FITC and incubated with chloroplasts. Binding reactions were performed in the presence of 30 μM ATP. Number of chloroplasts counted is shown on the y-axis and counts of S-Protein FITC labeling is shown on the x-axis. C. LSCM analysis of chloroplast binding of H-S-SStp constructs tagged with S-protein FITC. Chloroplasts are depicted in red and FITC signal is shown in green. D. Chemiluminescent detection of a Western Blot of a binding reaction showing interaction between H-S-SStp$_{WT}$ and H-S-SStp$_{\Delta 5}$ and chloroplasts. + fields show conditions in each lane. Binding reactions were performed in the presence of 30 μM ATP as described previously (Subramanian et al 2001). After the binding reaction, intact chloroplasts were subjected to SDS-PAGE. S-protein alkaline phosphatase was used for detection. Experiments performed by Chitra Subramanian.

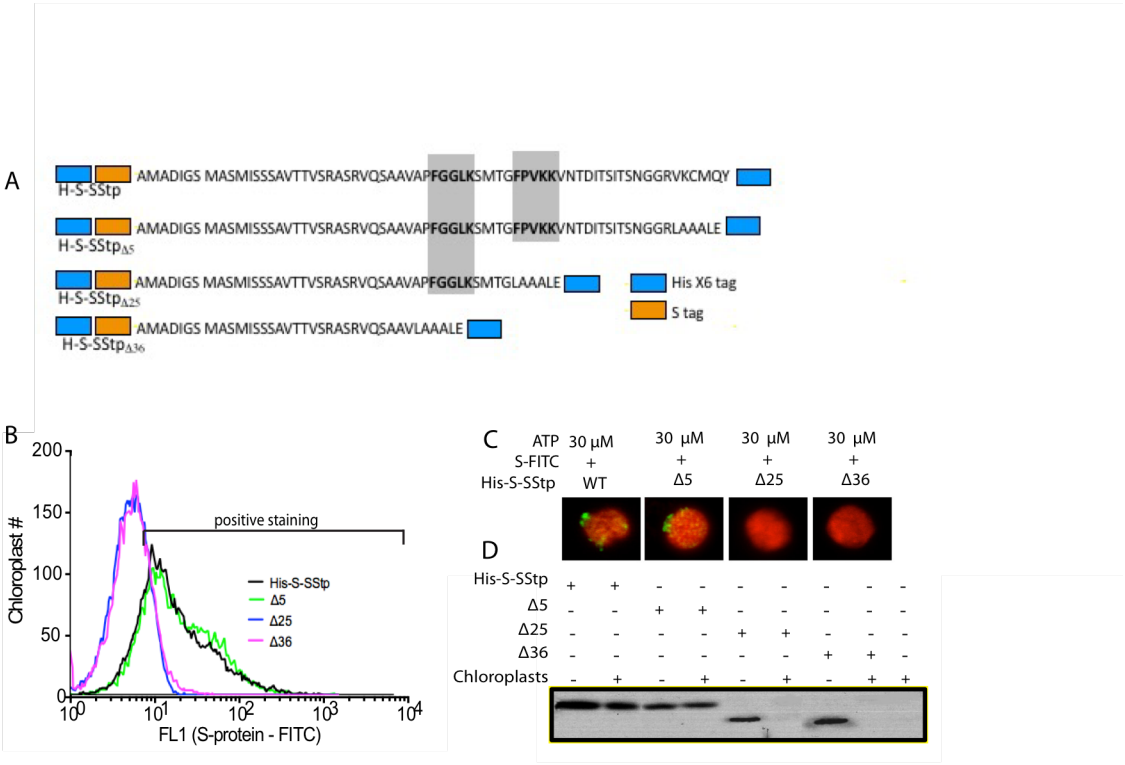


Figure 4-2 Continued

4.2.3 In vitro import competitions confirm the involvement of both FGLK domains

To specifically test the relative importance of the two FGLK domains in protein import, we have used *in organello* chloroplast import competition assays. We generated mutants of full-length tobacco RuBisCO precursor (prSSU) where we fully or partially deleted each of the two FGLK regions (denoted DI and DII). We constructed six deletion variants: ΔFTGLK (DIΔ5), ΔFTG (DI-NΔ3), ΔGLK (DI-CΔ3), ΔFPVSR (DIIΔ5), ΔFPV (DII-NΔ3), and ΔVSR (DII-CΔ3). The radiolabeled mutant precursor proteins, WT preprotein (prSSU), and mature protein (mSSU) were recombinantly expressed in *E. coli* and purified (data not shown**)**. They were then tested in chloroplast import competition assays. The concentration of ^{35}S-labeled prSSU was held constant at 100 nM while the concentration of the competitor protein was increased from 0 to 600 nM. Import of the labeled prSSU can be observed by quantifying accumulation of the radiolabeled mature domain in re-isolated chloroplasts. To quantify this competition ability we computed an IC_{50} value, defined as the concentration of inhibitor required for inhibition of 50% of import activity (Figure 4-3). We used the mature domain of small subunit RuBisCo (mSSU), which lacks the TP sequence, as a negative control unable to compete for import into chloroplasts ((Figure 4-4A-B).). Both the full DIΔ5 and DIIΔ5 deletion mutants showed a significant reduction in their ability to compete for import (Figure 4-4A) with an IC_{50} values more than tenfold higher than WT prSSU (Figure 4-3). All of the deletions are less active than the WT prSSU yet the deletion of either DI or DII raised the IC_{50} to >1700 nM. To test the relative importance of the N-terminal and C-terminal regions of the FGLK domain we also deleted the first and last 3 residues of each motif individually (Figure 4-4B).). Both of the N-terminal partial deletions have a less dramatic effect on competition ability raising the IC_{50} only to ~350 nM for DI-NΔ3 and with DII-NΔ3 somewhat higher with an IC_{50} of ~750 nM. The two C-terminal partial deletions, DI-C Δ3 and DII-CΔ3, both have substantially higher IC_{50} values values (~1000 nM) (Figure 4-3). Thus our data suggests that there is a both a specific importance for FGLK motifs and a requirement for the C-terminal residues (-GLK/-VSR).

Competitor	IC_{50} (nM)	Competitor	IC_{50} (nM)	Competitor	IC_{50} (nM)
WT prSSU	147 ± 20	DI-P 1A	85 ± 30	DI-F 1W	420 ± 90
mSSU	no inhibition	DI G 1A	130 ± 30	DII-F 1W	350 ± 70
ΔDI	1740 ± 330	DII-P 1A	120 ± 30	DI-F 1A	600 ± 200
ΔDII	1730 ± 210	DI-PG 2A	430 ± 20	DII-F 1A	660 ± 90
DI-N Δ3	340 ± 90	DI-P/DII-P 2A	130 ± 30	DI-F 1S	650 ± 130
DI-C Δ3	950 ± 180	DI-PG/DII-P 3A	no inhibition	DII-F 1S	730 ± 30
DII-N Δ3	750 ± 70			DI-F/DII-F 2W	310 ± 30
DII-C Δ3	1080 ± 270			DI-F/DII-F 2A	800 ± 150
				DI-F/DII-F 2S	1080 ± 200

Figure 4-3: IC_{50} values for *in organello* competition assays

Table lists IC_{50} values (the amount of competitor required for 50% inhibition of WT ^{35}S labeled protein into isolated chloroplasts) for competition assays

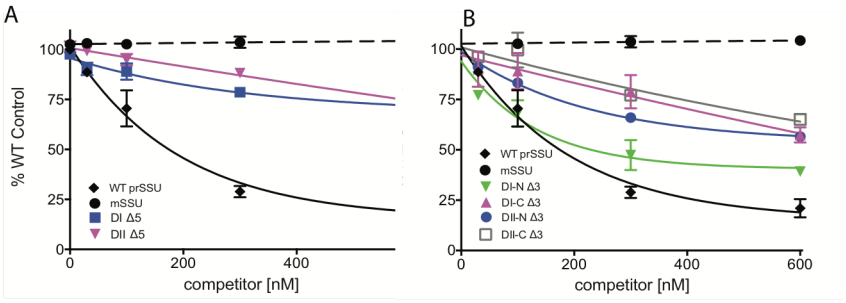


Figure 4-4: In vitro analysis of the SStp FGLK domain

A. Graphical analysis of *in vitro* competition assays using ^{35}S labeled prSSU with increasing unlabeled mutant competitor protein where the FGLK motifs (DI or DII) have been deleted. B. Graphical analysis of *in vitro* competition assays using ^{35}S labeled prSSU with increasing unlabeled mutant competitor protein with partial deletion of FGLK domains (N or C terminus of DI or DII). Experiments performed by Sarah Wright.

4.2.4 *In vivo* analysis suggests that multiple FGLK domains are redundant in function

Although our *in organello* data suggests a clear importance for each of the two FGLK motifs found in prSSU, the specific role of the FGLK motif *in vivo* remains poorly studied. *In vivo* analysis allows us to capture a long time scale, physiological response to compare with our short time scale, reductionist *in organello* biochemical assays. We have developed an *in vivo* assay, described in Chotewutmontri *et al*, using a SStp-20 mature domain residues-YFP chimeric construct to track the localization of mutant TPs in onion cells (Chotewutmontri and Bruce, 2015; Chotewutmontri et al., 2012). Although onion leucoplasts are non-green plastids, there are several recent *in vitro* and *in vivo* studies that show import to non-green plastids is similar to differentiated chloroplasts (Chu and Li, 2015; Hirohashi et al., 2001; Primavesi et al., 2007). Collectively, these works show that although there may be different expression levels of certain Toc GTPases and even possibly different import activities of these different plastid sub-types, there is a common mechanism of transit peptide recognition and translocation.

We generated a series of internal deletion constructs, removing either part or all of each FGLK motif. As a control against TP length-dependent effects (Bionda, 2010), we also constructed a series of Ser-substitution mutants, shown as green bars (Figure 4-5A). We selected Ser substitution because it is the most abundant amino acid in TP sequences (von Heijne et al., 1989b) suggesting a relatively neutral impact on TP structure and/or function. Transient expression of TP-YFP constructs and calculating a fluorescence ratio (plastid:cytosol) allows us to indirectly determine levels of YFP-protein import to the plastid compared to the surrounding cytosol; computing a ratio helps control for cell-to-cell variations in expression levels. When mutant constructs are not correctly targeted to the plastid, we observe YFP fluorescence in the cytosol and other cellular compartments including the nucleus (Figure 4-6).

The ratio of plastid:cytosol YFP signal intensity is reduced by 5% for the DIΔ5 construct; and 29% for the DIIΔ5 construct, which was found to be not statistically significant (quantification shown in Figure 4-5A; representative images shown in Figure 4-5B). However, deletion of both motifs in the same construct (construct DI/DIIΔ10)

Figure 4-5: *in vivo* localization assays reveal that semi-conserved motifs have a redundant role in transit peptide import

A. Transient expression of chimeric SStpNt-YFP constructs in onion (*Allium cepa*) were live-imaged using epi-fluoresecence microscopy. Alignment depicts location of mutations in the SStp sequence; domains DI and DII are highlighted in grey. Bars show the ratio of plastid localized signal intensity/cytosolic localized signal intensity. Blue bars denote deletion constructs and green bars denote Ser substitution constructs. "*" indicate cells where WT SStpNt-YFP displays a significantly higher Plastid/Cytosol ratio (more plastid labeling) than cells where SStpNt-YFP is more localized in the cytosol. (N=at least 15 cells per condition) One Way ANOVA with Bonferroni Post Hoc Analysis. DNI = Does not import; No stars = $P > 0.05$; * = $P \leq 0.05$; **= $P \leq 0.01$; ***= $P \leq 0.001$; **** = $P \leq 0.0001$ B. Representative epi-fluorescence microscopy images of the quantified data shown in A.

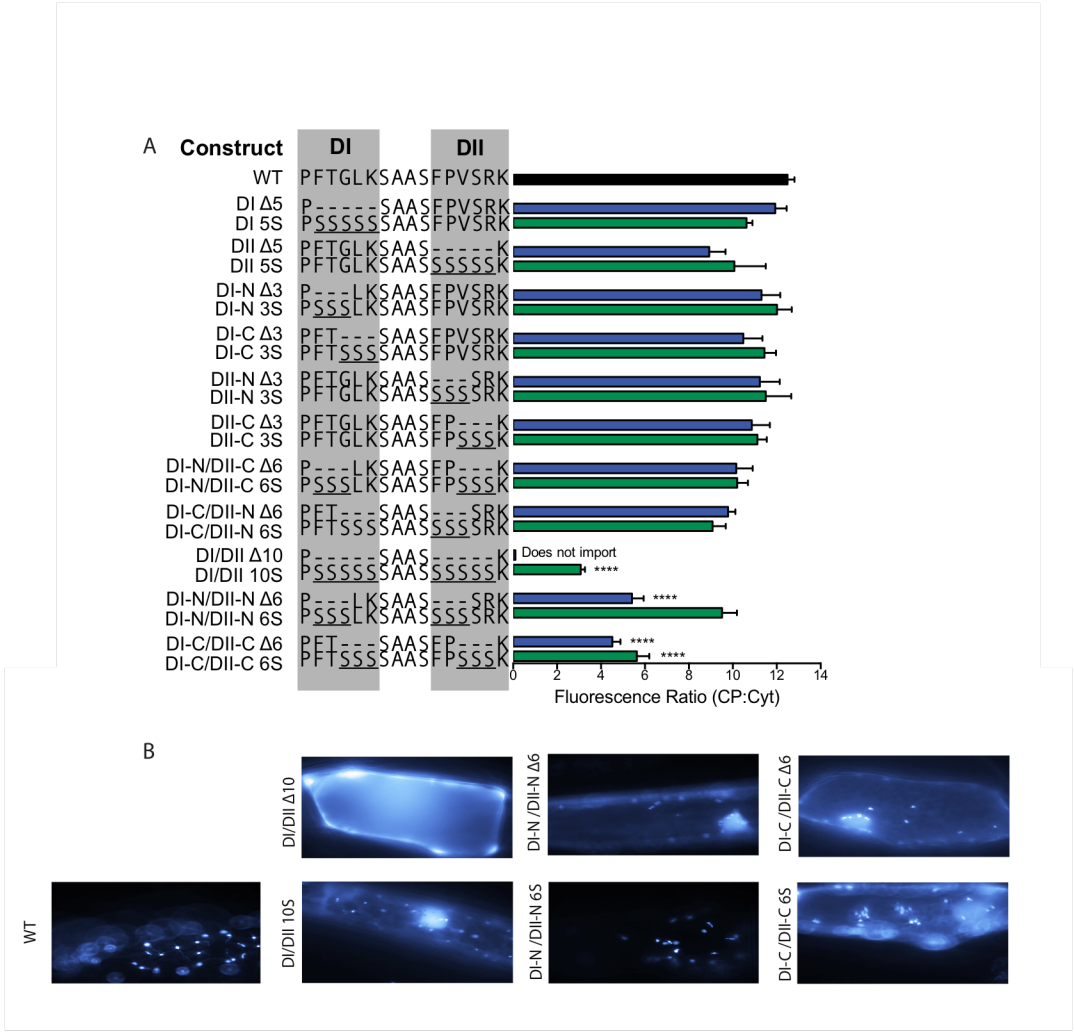


Figure 4-5 Continued

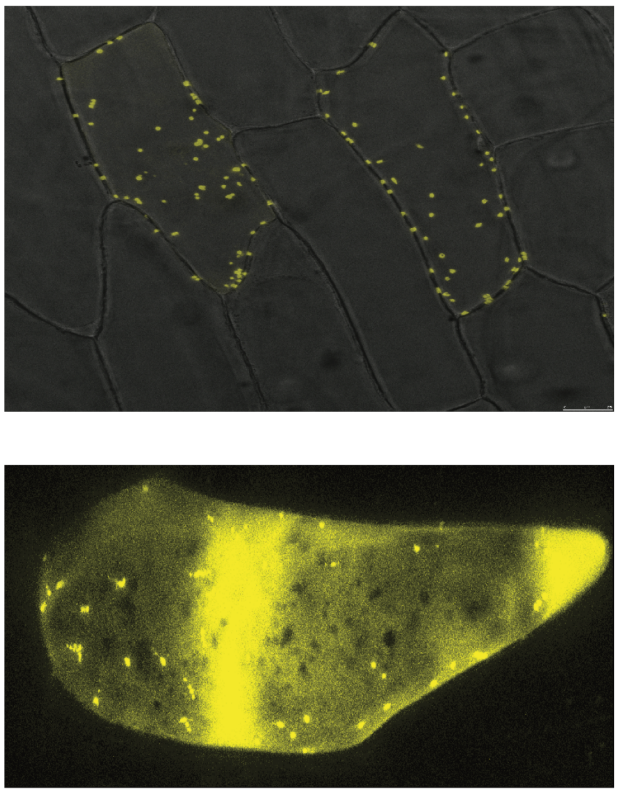


Figure 4-6: Confocal images of leucoplasts transformed with TP-YFP constructs

Z-series max projection of onion epidermal slices transformed with SStp-NT-YFP (WT) (Top Panel) or DI-5S/DII-5S-YFP (Bottom Panel). Images were taken on a Leica LSCM SP8 LightGate Microscope.

results in abrogated import (Figure 4-5A-B). Furthermore, Ser substitution of both of DI and DII motifs (construct DI/DII 10S) can not restore plastid localization to WT levels (Figure 4-5A-B), yet provides an increase to 23% of WT plastid:cytosol signal intensity. Interestingly, removal of only the C-terminal elements of both of the motifs (construct DI-C/DII-C Δ6) reduces import to 34% of WT levels, even when the region is substituted with Ser residues, import is only 43% of WT (construct DI-C/DII-C 6S) (Figure 4-5A-B). This data is in agreement with our import competition results shown in Figure 4-4. Deletion of the N-terminal elements (construct DI-N/DII-NΔ6) also results in cytosolic mislocalization of the protein; however, substitution with Ser (construct DI-N/DII-N 6S) restores to 73% of WT import levels. In addition, in every case Ser replacement of the all of the DI and DII double mutants either increased the import activity (DI/DII 10S; DI-N/DII-N 6S; and DI-C/DII-C 6S) and did not change it significantly from the deletions (DI-N/DII-C 6S; DI-C/DII-N 6S), implying that spacing may contribute function of these two domains *in vivo.*

4.2.5 *In vitro* analysis of the roles of aromatic residues in the FGLK domain

The first component of this motif is the inclusion of an aromatic amino acid such as Phe/Trp. We tested the role of the two Phe residues in Domain I/II motifs using *in organello* assays with isolated chloroplasts as described in Figure 4-4. We focused on generating several constructs in which both of the FGLK Phe are mutated to Trp, Ser, or Ala. These were produced via recombinant expression system in *E. coli* (data not shown). In import competition assays (Figure. 4-7A-D; Figure 4-3), mutation of aromatic Phe residues to aromatic Trp results in near-WT competition with ^{35}S labeled WT prSSU (Figure 4-7C), as expected. However, mutation to Ala (Figure 4-7A) or Ser (Figure 4-7B) significantly reduces competition (Figure 4-7D). This data suggests that loss of aromatic residues results in decreased TP recognition or import. In *in vivo* localization assays, we find that mutation of DI and DII Phe residues to Ala results in a 24% reduction in plastid import (quantification shown Figure 4-8A, IC_{50} values shown in Figure 4-3).

4.2.6 Structural flexibility affects binding, import and processing of TPs

TPs have been previously described (von Heijne, 1991) as the “perfect random coil;” some part of this lack of structure is due to their relatively high content of the helix breaking residues Pro/Gly. Due to their placement within the FGLK motif, we investigated

how these specific residues of the FGLK motif using *in organello* assays with isolated chloroplasts. To test the role of structural flexibility, we mutated the Pro and Gly residues located near DI and DII (Pro_{25}, Gly_{28}, and Pro_{36}) to Ala. We used an *E. coli* recombinant expression system to generate ^{35}S-labeled WT and mutant proteins for import time course assays. We incubated ^{35}S-prSSU WT and mutant proteins with isolated chloroplasts and monitored the rate of import and processing to the radiolabeled mSSU species. This assay allowed us to compare the kinetics import for mutant proteins compared to WT prSSU. We find that all three single Pro/Gly mutants (DI-P 1A, DI-G 1A and DII-P 1A) are able to be imported at a reduced rate. The DI-G 1A mutant is only reduced by 25%, yet the double Pro mutations show a reduction of 50-60% of WT SStp (Figure 4-9A). A second confirmation of the import ability of these P/G-A mutants is their competition for import at near-WT levels (Figure 4-9B). However, when we tested double and triple Pro/Gly substitution (double mutants DI-PG 2A, DI-G/DII-P 2A, DI-P/DII-P 2A, and triple mutant DI-PG/DII-P 3A), the triple mutant shows only 10% of the WT plastid localization (Figure 4-9C). In import competition assays, the triple mutant DI-PG/DII-P 3A is virtually unable to compete for import and the double mutants display an intermediate effect, with DI-PG 2A much more strongly affected (Figure 4-9D; Figure 4-3).

Because we established the importance of helix-breaking residues in *in vitro* analysis, we decided to test the *in vivo* localization of our Pro deficient mutants. We tested double substitution mutants DI-PG 2S and DI-PG 2A, and our triple mutant DI-PG/DII-P 3A; yet only the triple mutant DI-PG/DII-P 3A shows significant 39% reduction in plastid localization (quantification is shown in Figure 4-8B). The loss of plastid localization *in vivo* of the triple mutant agrees with the effects seen in Figure 4-9A-D, yet the observation that this *in vivo* effect is milder than the *in vitro* results is interesting. Removing the helix breaking residues P/G may extend the alpha helical nature of the TP. We have already shown that SStp has forms two helical segments when placed in a membrane mimetic environment (Bruce, 1998b). Increasing the helical length or stability of the TP could lead to a higher energy requirement for protein translocation. Since it is well established that protein import is driven by stromal ATPase activity (Liu et al., 2014; Shi and Theg, 2013; Theg et al., 1989b), we investigated if the import of these mutants could be increased by raising the levels of available ATP. However, even at very high ATP concentrations (>2

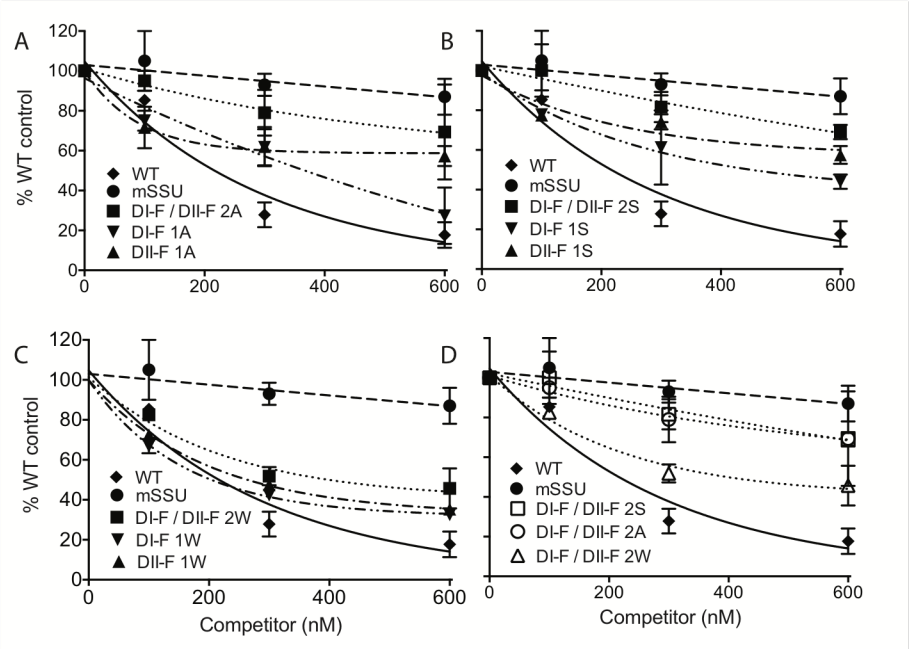


Figure 4-7: *in vitro* analysis of the role of aromatic residues in FGLK motif

A-C. Graphical analysis of *in vitro* import competition assays of Ala (A), Ser (B) and Trp (C) substitution mutants using ^{35}S labeled WT prSSU and increasing concentration of unlabeled mutants incubated with isolated chloroplasts. D. Graph summarizes the effects of DI/DII Phe substitution to Trp, Ser or Ala in *in vitro* competition assays using ^{35}S labeled WT prSSU and unlabeled mutants, as depicted in A-C. Experiments performed by Huixia Zhang.

Figure 4-8: Analysis of Phe and Pro mutations on TP import efficiency

A-B. Transient expression of chimeric SStpNt-YFP constructs in onion (*Allium cepa*) were live-imaged using epi-fluoresecence microscopy. Alignment depicts location of mutations in the SStp sequence; domains DI and DII are highlighted in grey. Bars show the ratio of plastid localized signal intensity/cytosolic localized signal intensity. Substitution to Ala is shown in gray bars/Ser is shown in green. N=at least 15 cells per condition. One Way ANOVA with Bonferroni Post Hoc Analysis. C. Graphical analysis of in vitro protein import assays using isolated chloroplasts and 35S labeled prSSU with increasing concentrations of [ATP], shown as a function of %WT mSSU import levels. [ATP] titration performed by Sarah Wright.

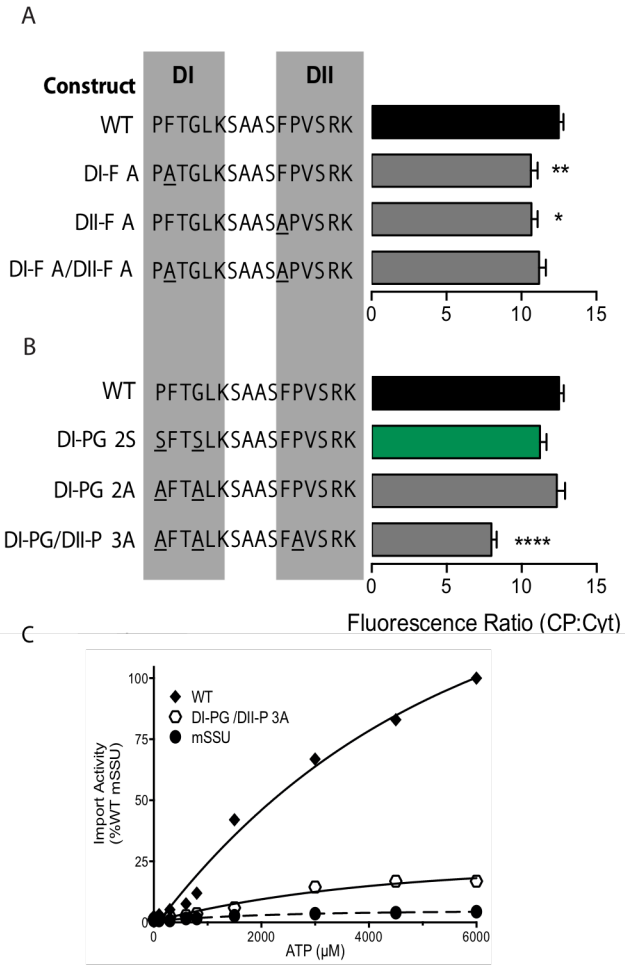


Figure 4-8 Continued

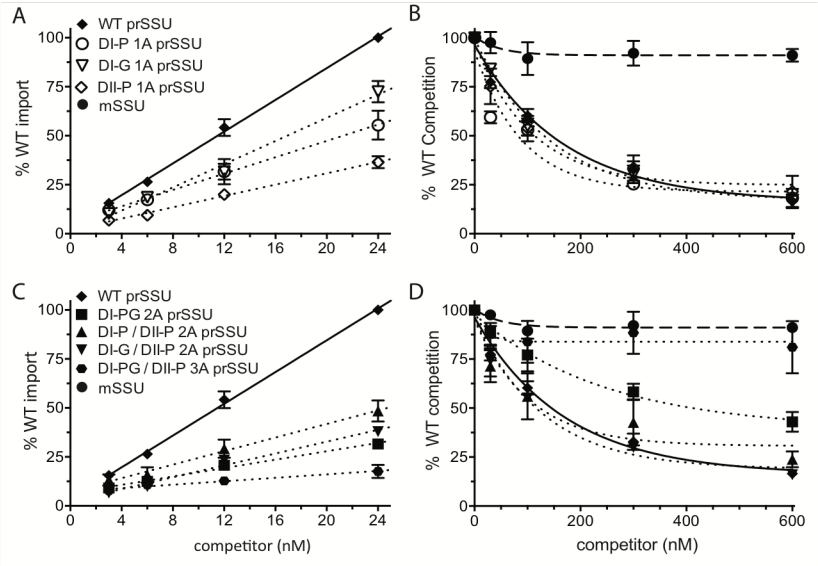


Figure 4-9: *in vitro* analysis of the role of aromatic residues in FGLK motif

A-C. Graphical analysis of *in vitro* import competition assays of Ala (A), Ser (B) and Trp (C) substitution mutants using ^{35}S labeled WT prSSU and increasing concentration of unlabeled mutants incubated with isolated chloroplasts. D. Graph summarizes the effects of DI/DII Phe substitution to Trp, Ser or Ala in *in vitro* competition assays using ^{35}S labeled WT prSSU and unlabeled mutants, as depicted in A-C. Experiments performed by Sarah Wright.

mM), the import machinery cannot efficiently transport the mutant precursor protein at WT levels (Figure 4-8C).

4.2.7 Loss of TP flexibility decreases interaction with psTOC34 *in vitro*

Both the *in vitro* and *in vivo* results suggest contributions of the complete Domains I/II, as well as roles for specific residues. In light of these results we investigated the effect of TP mutations on the interaction with the GTPase psTOC34 component of the TOC translocon, using *in vitro* assays (Chotewutmontri et al., 2012; Reddick et al., 2007b; Schleiff et al., 2002a; Sun et al., 2002a; Sveshnikova et al., 2000). We employed an *E. coli* based recombinant expression system to produce both TOC34 protein and full-length TPs, as previously described (Chotewutmontri et al., 2012; Reddick et al., 2007b). Two of these assays include the ability of WT SStp to induce monomerization of recombinant TOC34 and to function as GTPase Activating Partner (GAP)-like stimulatory activity on the GTP hydrolysis rate of psTOC34 (Chotewutmontri et al., 2012; Reddick et al., 2007b).

We used analytical ultracentrifugation (AUC) to monitor the oligomeric state of psTOC34 to test the overall contribution of the DI and DII domains of SStp, as well as the effect of helix breaking residues (Pro/Gly). We have previously shown that WT psTOC34 exists in a concentration dependent equilibrium between the monomeric and dimeric forms (Reddick et al., 2007b). When we add WT SStp from tobacco (Nt) or pea (Ps) the equilibrium strongly shifts psTOC34 toward a monomeric state (Figure 4-10A). However, incubation of psTOC34 with mutated TPs lacking either the DI domain, DII domain or the FGLK localized flexible residues (Pro_{25}, Gly_{28}, Pro_{36}) results in a reduced monomerization effect.

The disruption of homodimerization affinity was confirmed with a simple SDS-PAGE assay following covalent crosslinking using a sulhydryl-reactive, homobifunctional crosslinking agent, Bismaleimidohexane (BMH). BMH reacts with the single Cys residue (Cys_{215}) in psTOC34 that flanks the homodimeric interface observed in the crystal structure of TOC34 (Sun et al., 2002a). Figure 4-10B summarizes our crosslinking results with SStp and psTOC34. Under control conditions, addition of BMH stabilizes approximately 66% of the protein in a dimeric species on SDS-PAGE (Figure 4-10B, lane 2), yet addition of WT SStp protein strongly shifts the TOC34 protein toward a monomeric species, with only 15% dimeric species (Fig. 5B, lane 3). In agreement with our AUC

analysis, the triple mutant DI-PG/DII-P 3A does not monomerize TOC34 to the same extent as the WT SSTp protein, resulting in about 30% dimeric species (Figure 4-10B, lane 4).

We further tested the effect of the DI-PG/DII-P 3A substitution using a sensitive phosphate release assay to monitor psTOC34 GTPase activity (Chotewutmontri et al., 2012; Reddick et al., 2007b) (Figure 4-10C). PsTOC34 GTPase basal catalytic rates have been previously described (Reddick et al., 2007b); we find that titration of a 25 fold molar excess of WT SStp results in a maximal stimulatory effect, increasing the V_{max} by ~1.6 fold to 69.3 nmol GTP/min/µmol psTOC34. However, incubation with the same concentration of DI-PG/DII-P 3 mutant maintains near basal TOC34 GTP hydrolysis levels with a V_{max} of 32 nmol GTP/min/µmol psTOC34 and Km of 2.9 µmol. These two assays in combination indicate that the loss of flexibility of SStp directly reduces the productive interaction with Toc34 *in vitro.*

4.2.8 Basic residues required for *in vivo* import efficiency

Since our previous assays (Figure 4-4) suggested a unique requirement for the C-terminal portion of the DI and DII domains and since both of these domains contain at least one basic residue, we decided to more carefully test its importance by substituting the TP basic residues with either Ser or Ala (Figure 4-11A-B). We focused on *in vivo* analysis to test the importance of basic residues because our previous analysis established that only mutations with a strong *in organello* effect impact affect *in vivo* localization. Our previous *in vivo* localization analysis required substantial deletions or substitutions in DI and DII for construct mislocalization (Figure 4-5A-B). Mutation of a single basic residue (K_{30}, R_{39}, K_{40}) results in mildly reduced plastid localization levels (a 10% reduction for R_{39} and K_{40}); only K_{30} mutagenesis shows a significant 26% reduction (Figure 4-11A). However, we find that substitution of all three basic residues in the DI and DII domains results in a strong cytosolic mislocalization (Figure 4-11A). Substitution of K-RK with Ala results in a 78% reduction in plastid localization; Ser substitution reduces localization by 45%. Moreover, double mutations result in intermediate phenotypes (DII-RK 2S and 2A). There are two other basic residues in the SStp sequence, Arg_{14} and Arg_{54}, so we tested whether the mislocalization of DI-K/DII-RK 3A or 3S is due to loss of

Figure 4-10: The SStpNt DI-PG / DII-P 3A mutant shows reduced ability to interact with psTOC34 in *in vitro* experiments

A. Quantification of AUC analysis reveals that psTOC34 exists in a monomer-dimer equilibrium, and incubation with either SStpNt or SStpPs biases the equilibrium toward the monomeric form. The SStpNt DI-PG / DII-P 3A mutant does not significantly alter the equilibrium. B. Sulfhydryl-based crosslinking of TOC34 with TP mutants and nucleotides. + fields show reaction conditions for each lane. Monomer and dimer were separated using SDS-PAGE and quantified as a percentage of normalized dimer (basal conditions with BMH were normalized to 100%). WT and mutant transit peptide and were incubated with psTOC34 and crosslinked using the sulfhydryl-reactive crosslinking agent BMH. C. 25 fold molar excess WT SStpNt transit peptide stimulates psTOC34 GTP hydrolysis in a radiolabeled γ^{32}P-GTP experiment, but the same concentration of SStpNt DI-PG / DII-P 3A mutant cannot. The γ^{32}P-GTP was held constant at 10nM throughout the experiment. Counts/min of TOC34 hydrolysis at increasing GTP concentration were used to calculate a slope that was mathematically transformed into a rate expressed as nanomoles of GTP hydrolyzed per min per µmol TOC34, and graphed to yield a Michaelis-Menten substrate velocity plot. Experiments were carried out essentially as described in (Reddick et al., 2007b). Experiments performed by Sarah Wright and Evan Reddick.

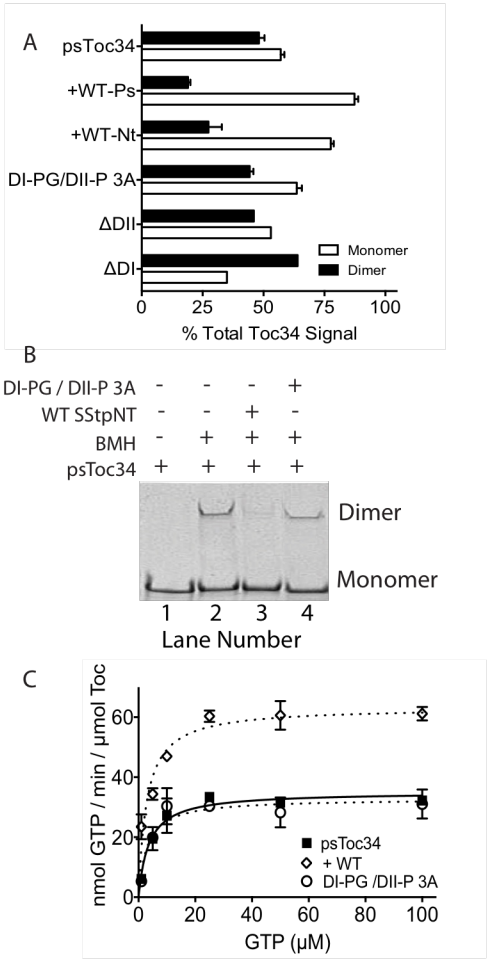


Figure 4-10 Continued

Figure 4-11: *in vivo* localization assays reveal that basic residues located in the FGLK domain are required for import efficiency

A. Transient expression of chimeric SStpNt-YFP constructs in onion (*Allium cepa*) were live-imaged using epi-fluorescence microscopy. Alignment depicts location of mutations in the SStp sequence; domains DI and DII are highlighted in grey. Bars show the ratio of plastid localized signal intensity/cytosolic localized signal intensity. Blue bars denote deletion constructs and green bars denote Ser substitution constructs; substitution to Ala is shown in gray. (N=at least 15 cells per condition) One Way ANOVA with Bonferroni Post Hoc Analysis B. Representative epi-fluorescent microscopy images of Ala substitution mutants. Image quantification is depicted in A.

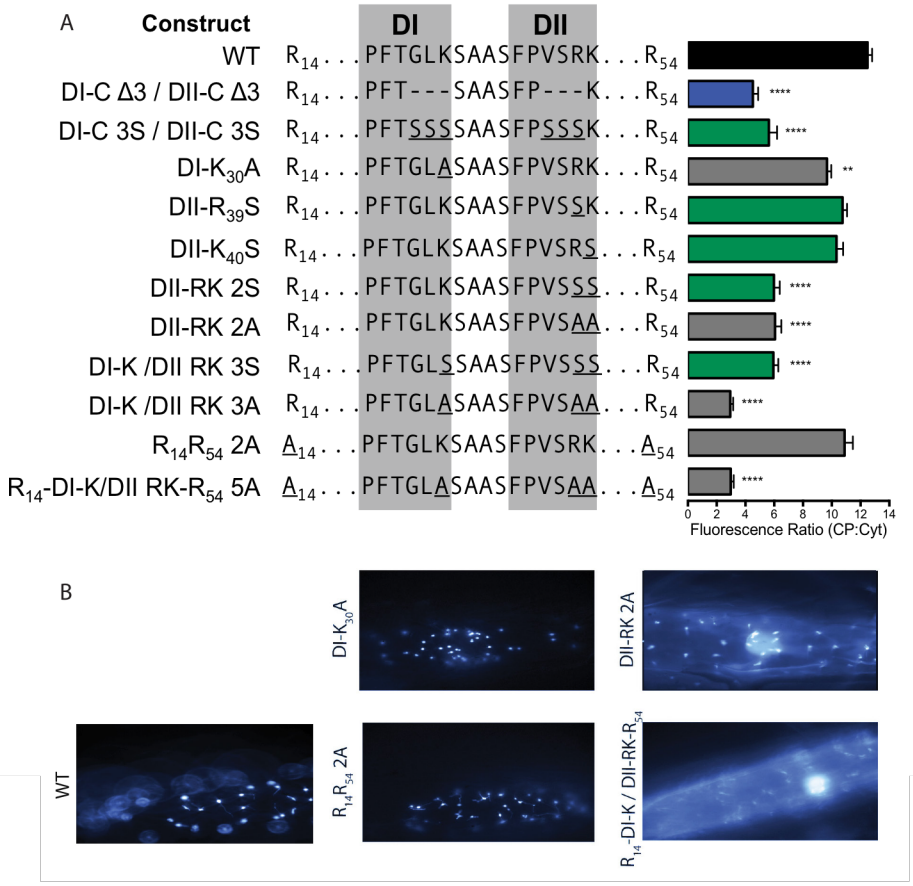


Figure 4-11 Continued

basic residues in the physiochemical DI/DII domains, or if the effect is nonspecific to loss of positive charge. We find that substitution mutant $R_{14}R_{54}$ 2A has no effect on protein localization compared to the WT control protein. Furthermore, addition of $R_{14}R_{54}$ mutation to the DI-K/DII-RK 3A construct (R_{14}-DI-K/DII-RK-R_{54}) does not increase mislocalization of the DI-K/DII-RK 3A construct (Figure 4-11A-B).

4.2.9 Importance of FGLK domains other TP sequences

Since our analyses suggest that FGLK is a required element in TP import in SStp, we decided to extend our analysis to a second model TP. The ferredoxin TP (FdTP) is probably the second most well characterized transit peptides (Chotewutmontri et al., 2012; Pilon et al., 1992a; Pilon et al., 1992c; Pilon et al., 1995b). Both of these are highly abundant preproteins targeted to the stoma of chloroplasts. Much of the early work focused on the FdTP from *Silene latifolia* (*S. pratenis*)(Smeekens et al., 1985). Alignment of ~40 FdTPs from angiosperms (Figure 4-12) indicate that these TPs also contain two FGLK domains (labeled as Domain I and Domain II) with a consensus sequence FLRKQP and FGLK, respectively.

Based on this alignment we identified two regions in the *Silene* precursor that contain the basic elements of the FGLK motif, denoted as DI-DII (Figure 4-12 and Figure 4-13). We systematically mutated these regions to Ser and used our *in vivo* YFP chimera assay to test the localization of the WT versus mutant constructs. We find that mutation of DII alone (FGLK; DII-4S) has no effect on the plastid localization of FDtp; however, mutation of DI (PKQQPM) in combination with the DII domain (DI-6S/DII-4S/RRS) abrogates chloroplast localization, similar to what we observed with SStp. To test the relative contributions of the amino acids in DI, we performed a similar Ser replacement tested the effect of substitution of the N- (PKQ) and C- (QPM) regions. We find that DI-N substitution (DI-N 3S) results in about a 29% reduction in import; however mutagenesis of DI-C (DI-C 3S) results in a 41% reduction from WT localization. However, when DI-N/DI-C is mutated in combination with DII, the construct is significantly mislocalized, indicating the additive effect of DI and DII domains and suggesting that these two FGLK motifs are redundant in function.

Figure 4-12: Alignment of the Ferredoxin Transit Peptide.

Angiosperm transit peptides have been aligned using the NCBI Blast Alignment tool with manual adjustment to connect the two “FGLK” domains. Amino acids within the transit peptide have been colored in the following ways: basic residues (R/K/H), helix breaking residues (G/P), Aromatic residues (F/W/Y), polar residues (N/Q) and acidic residues (D/E). The potential phosphorylated serine and threonine residues located between Domain I and Domain II are show as S/T. The sequences are listed alphabetically with the *Silene* sequence placed above the alignment and the consensus sequence at the bottom. The predicted SPP cleavage site is shown with the vertical line and the two FGLK domains are shown with a shaded box and a box within the consensus. A third partially conserved domain is shown with a dashed box in the bottom consensus.

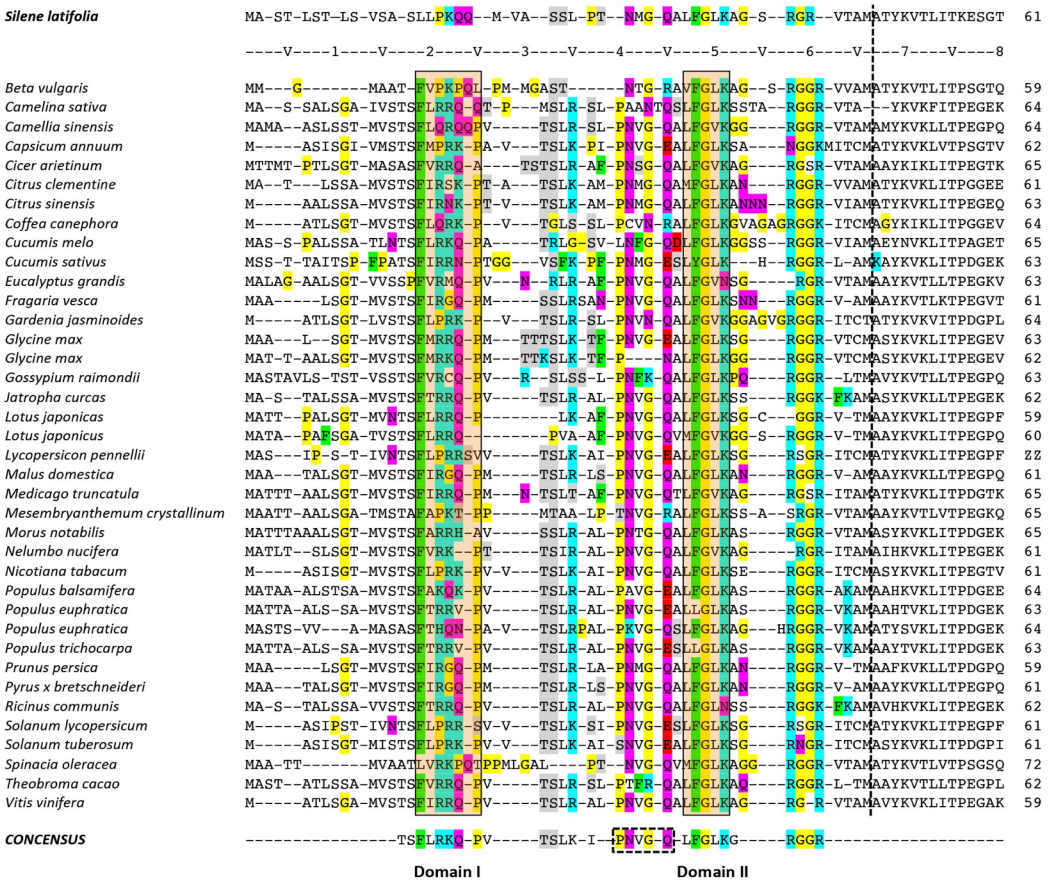


Figure 4-12 Continued

Figure 4-13: *In vivo* localization reveals FGLK domains in FdTP are required for efficient plastid localization

A. Transient expression of chimeric FDF-YFP constructs in onion (*Allium cepa*) were live-imaged using epi-fluorescence microscopy. Alignment depicts location of mutations in the FDF sequence; domains DI and DII are highlighted in grey. Bars show the ratio of plastid localized signal intensity/cytosolic localized signal intensity. (N=at least 15 cells per condition) One Way ANOVA with Bonferroni Post Hoc Analysis B. Representative epi-fluorescence microscopy images of the quantified data shown in A.

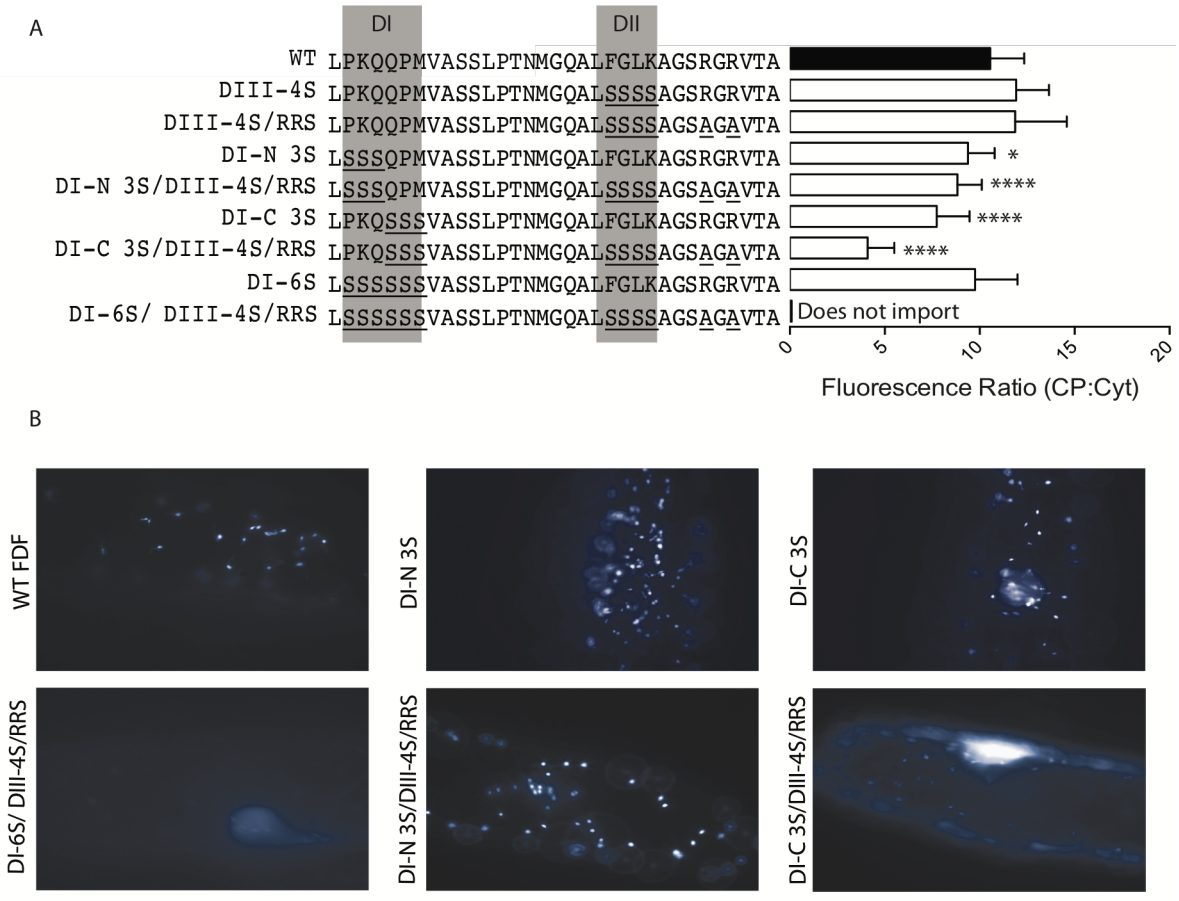


Figure 4-13 Continued

4.2.10 TP phosphorylation is not required for plastid protein import

Earlier work has shown that chloroplast precursors can be phosphorylated at specific Ser (Martin et al., 2006; Waegemann and Soll, 1996). These Ser residues have been partially mapped in prSSU to a motif that falls between the two FGLK motifs. The result of the phosphorylation has been proposed to enhance the affinity of the TP for the 14:3:3 proteins found within a guidance complex (May and Soll, 2000a). Because of this activity and the proximal location of this phosphorylation site, we decided to directly test the role of phosphorylation in precursor recognition and import into leucoplasts using a series of non-phosphorylatable or phosphomimetic mutations in FDtp and SStp (Fig. 8A-B). Analysis of the ~40 angiosperm ferredoxin transit peptides indicate that there is a variable number (from 2-5) of phosphorylatable residues between Domain I and Domain II (Figure 4-12). In the *Silene* Fdtp sequence, we selected 3 possible phosphorylation sites in the FDtp sequence, S_{24}, S_{25} and T_{28} (denoted SST). We then made pairwise comparisons of the non-phosphorylatable S-to-A mutants to the WT, as well as a pairwise comparison between non-phosphorylatable to the phosphomimetic mutants in FDtp (Figure 4-14). In SStp, we selected residues that have been previously implicated as phosphorylation sites (S_{31}/S_{34}, T_{27}/S_{38}) (Figure 4-14). Interestingly, in each of the non-phosphorylatable mutations (S/T to A) the import was actually slightly higher than the WT FDtp. This ranges from ~6% in ASA to a high of 33% in SAA, and overall these averaged about 22% above the WT FdTP; in no case was the introduction of a non-phosphorylatable Ala residue statistically lower than the WT FdTP. We observed a similar slight stimulation with the two Ser residues in SStp when they were changed to Ala.

However, upon mutation of the Ser to a phosphomimetic reside (S/T to D/E) we observed a general decrease in import. This ranged from a minor but not statistically significant change in T_{28} (SSA vs. SSE) to a S_{25}-to-D (SDT) compared to very strong effect for S_{24} (DST) which resulted in a 30% reduction in correct localization (Figure 4-14). The change in T_{28} was intermediate with about a 20% reduction in plastid accumulation. This effect was somewhat additive since we find that the double substitution SDE and DDT mutants show about 35% lower plastid import vs the correlating single Ala mutants (SAA and AAT, respectively). Similar to FDtp, SStp S_{31}/S_{34}A also shows an increase in plastid localization, whereas T_{27}/S_{37} results in about

a 30% reduction in protein localization (Figure 4-14). Overall, we find that the replacement of S/T to a non-phosphorylatable alanine. consistently result in plastid localization equal to or greater than the WT counterparts. However, these results indicate that the introduction of the phosphomimetic residue, Asp, consistently lowers the targeting activity *in vivo* in both SStp and Fdtp.

4.2.11 Potential Role for TP *cis*-Prolines

We hypothesize that specific *cis* Xaa-Pro residue(s) could alter the structure of the TP to increase or decrease recognition with one or more of the TOC components and affect precursor import efficiency. To further investigate the predicted isomeric form of Pro residues in the vicinity of the FGLK domains, we used CISPEPpred prediction software to predict Pro *cis/trans* peptide bond conformation in the RuBisCo small subunit TP (Figure 4-15 A & B) (Song et al., 2006)). We aligned 24 SStp sequences from green plants using ClustalX and overlayed the CISPEPpred prediction for *cis* Xaa-Pro residues. Although only three Pro positions were largely conserved (near positions 22, 30, & 41) only the Pro at position 30 was predicted to have a potential *cis* confirmation (Figure 4-15 A & B). This residue was predicted to be a *cis* Xaa-Pro in approximately 70% of the TPs. Interestingly, all of the *cis* Xaa-Pro residues are predicted to occur in the middle of the SStp sequence.

Next, we wanted to investigate the occurrence of *cis* Xaa-Pro residues in a database of 912 confidently predicted TPs using CISPEPpred (Chotewutmontri et al., 2012). Interestingly, this program predicted that ~35% of the TPs contained at least one *cis* Xaa-Pro. This is significant since only ~5% of all Pro are found in the *cis* confirmation of proteins with known structures (Jabs et al., 1999; Pall and Chakraabarti, 1999; Weiss et al., 1998). Recent studies have suggested that *cis* Pro conformation correlates with local amino acid sequence content (Song et al., 2006). When we looked at the flanking sequences (+/- five residues, xxxxxPxxxxx) of the predicted *cis* (Figure 4-15 C) and predicted *trans* (Figure 4-15D) Pro sequences from the Arabidopsis 912-TP database using a logo plot, we saw a clear difference in the frequency of flanking amino acids. The *cis* Pro residues have a very high occurrence of additional Pro residues that seem to be found on both sides of the *cis* Pro. We also see higher occurrence of Phe residues immediately flanking the *cis* Pro. However, the *trans* Pro residues seem to have a very

uniformly distributed pattern with S>L>P>R/K in the logo plot, occurring nearly equally on both sides of the *trans* Pro. However, it is difficult to discern the causality of these differing amino acid distributions.

4.3 Discussion

Despite the wide divergence of TP sequence and structure (Bruce, 2000b), previous studies implicate two physicochemical motifs in precursor protein targeting and import (Chotewutmontri et al., 2012; Ivey et al., 2000b; Lee, 2006, 2009; Pilon et al., 1995b). One motif is an uncharged N-terminal HSP70 interaction domain that functions to engage stromal ATP-dependent chaperones driving translocation (Chotewutmontri & Bruce, 2015). Considerable evidence now suggests HSP70 functions as this translocation motor for a subset of precursors (Chotewutmontri et al., 2012; Liu et al., 2014; Shi and Theg, 2010a). We have previously shown that this motif is followed by one or more FGLK motifs, which were initially reported many years ago, yet detailed description of their function is still lacking (Pilon et al., 1995b). Importantly, FGLK motifs represent a physicochemical element versus a sequence consensus. In this work, we greatly extend the analysis by combining *in vitro*, *in organello,* and *in vivo* assays. This provides a highly quantitative and resolved biochemical insight while confirming a clear physiological role in living cells.

4.3.1 FGLK domains are required for *in organello* binding

In prior work, chimeric constructs placed the reporter protein C-terminal to the TP sequence, analogous to a native TP domain structure. These studies showed robust activity of the chloroplast translocator, resulting in complete translocation of the chimeric fusion protein into the stroma (America et al., 1994). However, our previous work (Subramanian et al., 2001) characterized H-S-SStp, which could not be translocated into the chloroplast stroma yet was able to compete with TP for preprotein binding to intact chloroplasts. The addition of the S-tag provides a sensitive and quantitative reporter via its interaction with S-protein (Kim and Raines, 1993). The chimeric construct places 21 charged residues at the N-terminus of SStp, which conflicts with a widely regarded property of TPs (Chotewutmontri & Bruce, 2015; (Chotewutmontri et al., 2012; Pilon et al., 1995b; Subramanian et al., 2001; von Heijne, 1991). Despite the inability of this construct to translocate across the chloroplast envelope

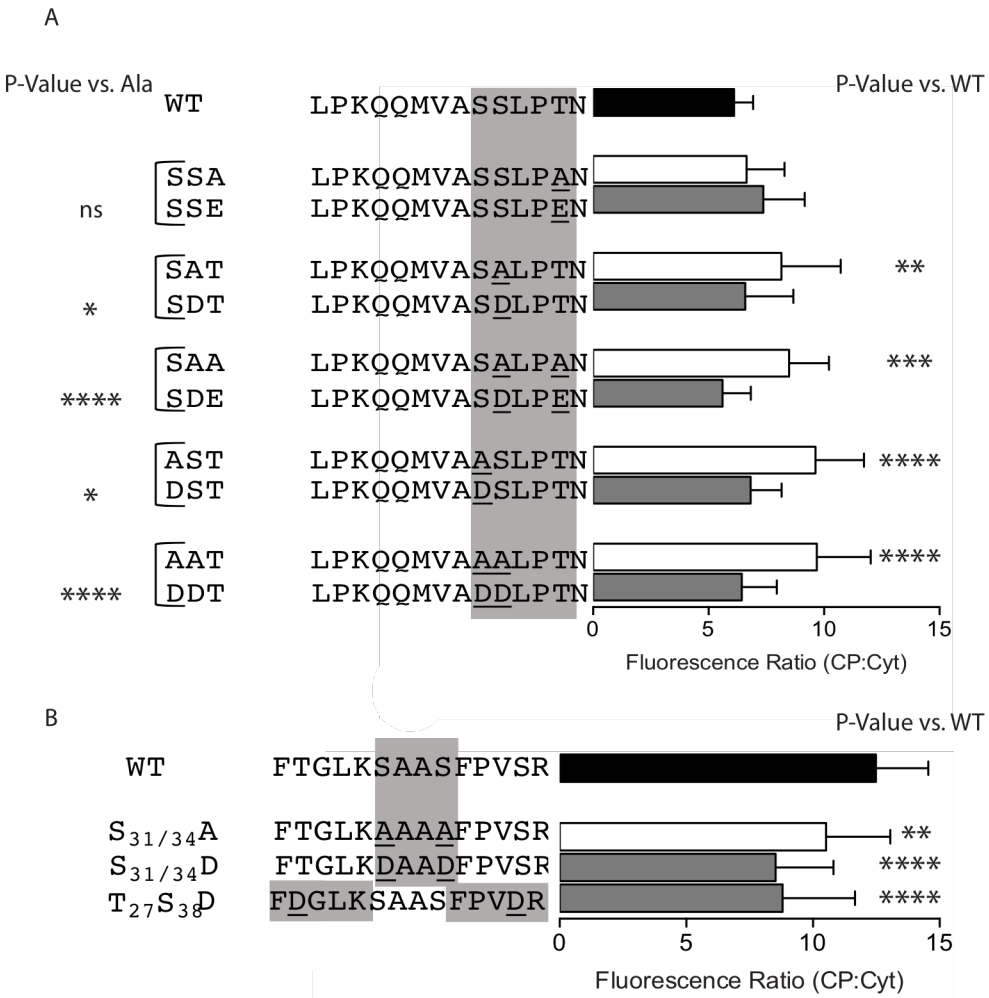


Figure 4-14: Phosphorylation of FDtp and SStp is not required for plastid import

A-B. Transient expression of chimeric FDF-YFP constructs (A) or SStpNT-YFP constructs (B) in onion (*Allium cepa*) were live-imaged using epi-fluorescence microscopy. Alignment depicts location of mutations in the FDF sequence. Bars show the ratio of plastid localized signal intensity/cytosolic localized signal intensity. (N=at least 15 cells per condition) One Way ANOVA with Bonferroni Post Hoc Analysis.

Figure 4-15: Potential Role of Cis-Proline in TP recognition

(A) Alignment of 24 RuBisCo small subunit TPs using ClustalX; Pro residues are highlighted in black. (B) Graph summarizes the locations of aligned Pro residues from (A). Cis-Pro were predicted using CISPEP-pred online tool at http://sunflower.kuicr.kyotou.ac.jp/~sjn/cispep/. (C-D) Logo plots of +5 and -5 amino acid positions around a cis-predicted Pro (C) and a trans-predicted Pro (D). Data set is 912 Arabidopsis TP sequences as identified in Chotewutmontri et al. (2012). Cis-Pro were predicted in the data set using CISPEP-pred. FDF analysis performed by Lily Moncrief.

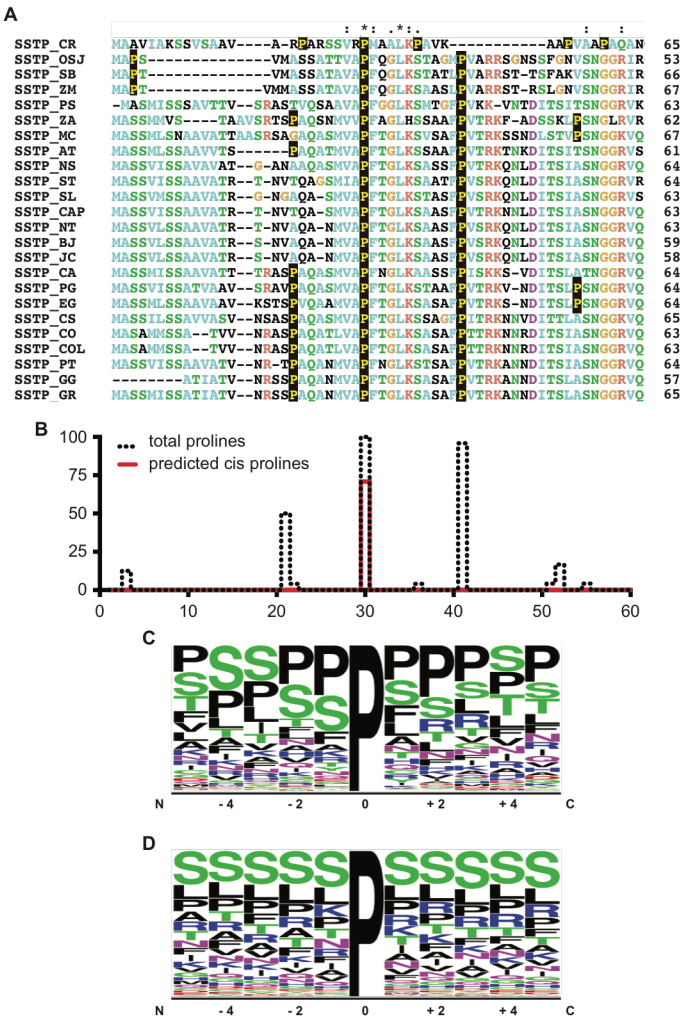


Figure 4-15 Continued

(Chotewutmontri et al., 2012; Subramanian et al., 2001), it is an ideal substrate to explore binding to the chloroplast surface.

Previous TEM microscopy showed that the inner and outer chloroplast membranes come together at punctate contact sites (Cline et al., 1985), and biochemical analysis of the TOC-TIC core complex suggests that membrane contact sites represent the locations of active protein transport (Schleiff et al., 2003b). Thus, we conclude that the punctate fluorescent pattern observed for the WT and Δ5 constructs represent H-S-SStp bound to the TOC-TIC supramolecular complexes (Figure 4-2C). Prior work has also demonstrated a punctate TP immunofluorescent imaging pattern (Schnell and Blobel, 1993a; Subramanian et al., 2001); however, unlike the divalent antibody based studies, our labeling approach uses a monovalent S-Protein FITC. This assay obviates the possibility that the punctate labeling arises from capping or clustering due to the multivalent nature of antibodies. It is not clear why they are punctate in nature but suggests an additional level of supramolecular organization. BN-PAGE has observed at least three sizes of the TOC complex so this punctate fluorescent labeling may be detecting these larger complexes (Chen and Li, 2007a; Kikuchi et al., 2006b).

The C-terminal TP truncations clearly disrupt H-S-SStp interaction with isolated chloroplasts (Figure 4-2A-D). It is clear from the differential effect of Δ5 vs. Δ25, that the critical residues are between G26 and R52. This region contains all of the second, C-terminal FGLK motif, which in pea is FPVKK. Interesingly, when we test the deletion of either the full domain I or II in import competition, both are nearly equal in their ability to compete with prSSU. This confirms that it is FGLK domains themselves that are key for recognition and binding to chloroplasts. However, it is not clear why the two assays show somewhat different importances of the two domains. The charged N-terminus of His-S-SStp possibly counteracts or reduces the efficacy of the single N-terminal FGLK domain to support binding. Interestingly, a basic residue doublet (RxRy) located within a FGLK domain has been implicated as a recognition sequence element necessary for TP recognition/import in older chloroplasts (Ling et al., 2012). Mutating this basic residue doublet in the *Tic40* precursor abrogated import of the precursor into older chloroplasts but not younger chloroplasts (Yi-Shan et al., 2012).

4.3.2 Differential role of from the individual amino acids

We further resolve this interaction by comparing the deletion of either the first three residues of either DI (FTG) or DII (FPV) versus the last three residues of DI (GLK) or DII (VSR), revealing a more severe result when the C-terminal residues were removed (Figure 4-4). However, both the N or C terminal mutants were more active than the complete removal of either DI or DII, suggesting that productive import involves all of physicochemical elements in the FGLK domain. Since the Ser substitution partially restores import of the DI-N/DII-NΔ6, it suggests spacing may be critical, but could also suggest that simply removing the N-terminal residues may eliminate some secondary structural elements that affect TP recognition. These results indicate that an element(s) in the C-terminal regions of the FGLK motifs, such as positively charged residues, may play a critical role in binding and/or import efficiency (Figure 4-11A-B).

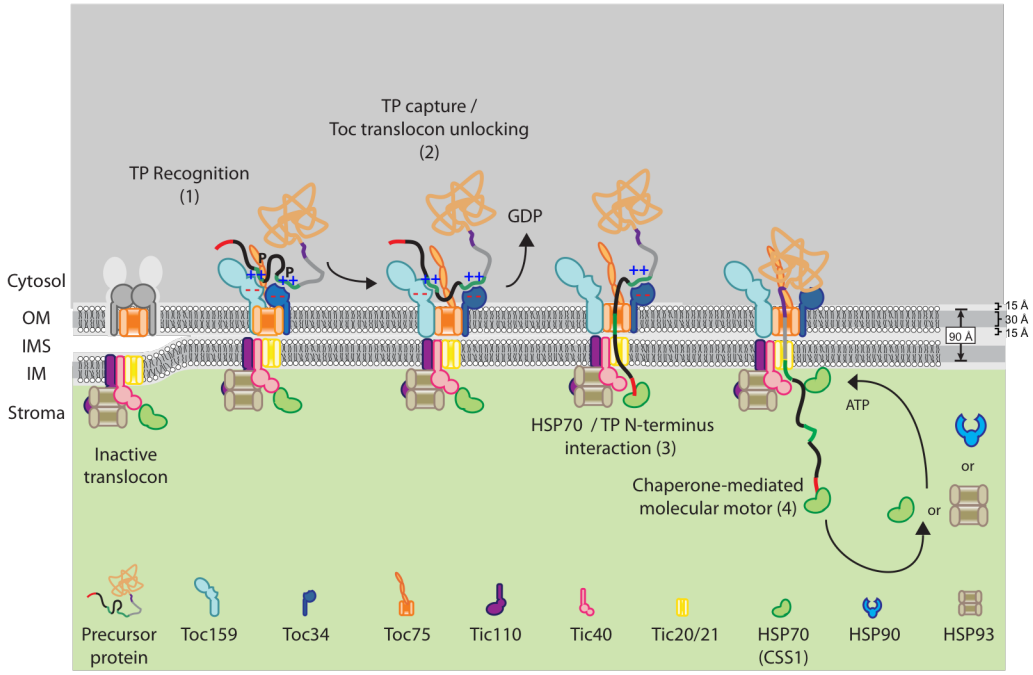


Figure 4-16: Bimodal model for TP import

We propose a dynamic recognition model, where (1) cytosolic TP recognition is driven through productive interaction between the TOC receptor components, such as TOC159 and TOC34. This is likely driven through electrostatic interactions involving Arg residues in the TP FGLK domains. Initial recognition may occur through interaction between TP hydrophobic residues and chloroplast lipids (not shown). (2) Following initial recognition, the TP is captured and the TOC translocon is made accessible, possibly through structural changes driven by GTP hydrolysis. (3) Cytosolic capture is followed by recognition of sequence elements in the N-terminus of the TP and stromal chaperones. (4) “Pulling” of the TP into the chloroplast by stromal chaperones such as chloroplastic HSP70 and HSP93.

Furthermore, import analysis of Arabidopsis E1αtp (the TP of E1α subunit of pyruvate dehydrogenase) found that both block Ala mutagenesis of a FGLK domain or a region containing 3 basic residues resulted in significant equivalent reduction in import (Lee, 2009). In agreement, our data suggests a dose-dependent threshold for basic residues in import efficiency; loss of 2 or more basic residues within the FGLK domains abrogates import (Figure 4-11A). The functional contribution of the positive charge requires their centralized placement proximal to the FGLK recognition elements, since removal of flanking residues at the N- and C-terminus (R_{14} and R_{54} respectively) had no effect on import efficiency in SStp (Figure 4-11A). Thus it appears that a direct correlation exists between the basic residues in the FGLK domain and TP recognition/import. Future work will explore the structural basis of this interaction.

The effects of mutation on FGLK components such as Phe (Figure 4-7; Figure 4-8A&B; Figure 4-9) are milder in *in* vivo experiments versus *in* vitro. Collectively, our data sugessts that there may be factors present in the cytsol and absent in our *in organello* assays that contribute to the import process, which negate the effect of Phe or Pro/Gly substition on TP secondary structure. Furthermore, this implies that the basic residues have a greater contribution to TP recognition/import since their mutagenesis abrogates import *in vivo* (Figure 4-11A-B).

We further verified the redundancy of multiple FGLK domains using *in vivo* localization analysis of FdTP, which also contains 2 FGLK physicochemical motifis. Similar to SStp, mutagenesis of only one FGLK domain resulted in WT plastid localization; whereas mutagenesis of both domains abrogated import (Figure 4-13). Collectively, these results suggest that the *in vivo* import can utilize either DI or DII with some redundancy. It is also interesting that partial deletions of FGLK domains in SStp still show some ability to compete for preprotein recognition (Figure 4-4), suggesting that the two domains may be somewhat redundant in promoting recognition. How these two domains are recognized is not known in detail, but it is possible each element is recognized by adjacent monomers in either a hetero- or homo-dimeric form of the TOC receptors, TOC34 and/or TOC159 (Lumme et al., 2014).

4.3.3 Direct Effects on TOC34 Structure and Activity

The effects of TP-TOC34 interaction on Toc34 homodimeric state and basal GTP hydrolysis activity has potential implications in the regulation of chloroplast protein import. Previously, the GTPase cycle of Toc34 has been hypothesized to function as a major regulatory switch in protein import. There is considerable evidence that the homo- and possible heterodimerization of Toc34 contributes to its GTPase cycle (Aronsson et al., 2010; Jelic et al., 2002a; Lumme et al., 2014; Reddick et al., 2008a). Previous experimental data supports that the FGLK motif functions as a TOC34 binding element (Chotewutmontri et al., 2012; Lee, 2006; Reddick et al., 2007b; Sun et al., 2002a). In the future, it would be interesting to see if the FGLK sequence can interact directly with TOC34 and either disrupt the stability of the dimer or stimulate GTPase activity. We find that removal of Pro residues from the vicinity of the FGLK motifs inhibits productive interaction between SStp and TOC34, both in AUC analyses and GTPase-stimulating experiments (Figure 4-10A-C). Although no high resolution structural information is currently available for the TP-TOC34 complex (likely due to low affinity of the complex), we suggest that loss of Pro residues near the FGLK motif disrupt secondary structural elements that are induced upon TOC34 interaction. Arg finger motifs have been identified in GTPase activating proteins (GAPs) in different orientation, suggesting that Arg fingers rotate into a catalytic position at the transition state (Sprang, 1997). It is plausible to hypothesize that loss of TP Pro residues prevents productive TP Arg orientation, thus inhibiting the productive interaction between TOC and TP. Interestingly, even at high ATP concentrations; the import machinery cannot overcome the loss of Pro residues in the FGLK domain (Figure 4-8C). Thus, it is tempting to speculate that the (R/K) basic residues and their orientation in the FGLK motif might also interact with TOC34 in such a way that it also stimulates GTP hydrolysis, which would bridge the importance of the FGLK domain with the proposed GAP function of the TP. Alternatively, the loss of flexibility of the TP may hinder the probability or kinetics of TP insertion into the protein conducting channel and/or diminish the TP N-terminus interaction with the stromal ATP-dependent molecular motors (Chotewutmontri and Bruce, 2015). Furthermore, it is possible that the removal of these prolines may preclude TP from inducing conformational changes in the Toc34. This may prevent changes in the homo- or heterodimer that would normally lead to TP

insertion into the translocon. In summary, it is clear that specific components of the FGLK domain have a direct role on TP recognition by Toc34 or other Toc receptor(s).

4.3.4 Role of TP phosphorylation in plastid import

Previous reports have hypothesized that a cytosolic kinase phosphorylates Ser and Thr residues in the TP consensus motif (P/G)Xn(R/K)Xn(S/T)Xn(S*/T*) were n = 0-3 amino acids spacer and S*/T* represents the phosphate acceptor (Waegemann and Soll, 1996). However, the detailed role of TP phosphorylation in this process is still remains poorly characterized and there are conflicting results about the importance of phosphorylation in TP recognition and import (Chen et al., 2014; Lamberti et al., 2011b; Nakrieko et al., 2004; Nickel et al., 2015). Because this proposed consensus motif can overlap with TP FGLK motifs, we decided to analyze the effect of phosphorylation in two different TPs using our *in vivo* protein localization assay. To test the importance of TP phosphorylation in plastid import, we designed mutations to either mimic the negative charge associated with a phosphorylated Ser/Thr or replace the Ser/Thr with a non-phosphorylatable residue, alanine. We find that non-phosphorylatable mutants consistently show a stronger plastid localization versus their phosphomimetic counterparts (Figure 4-14). In support of our data, previously, SStp S31/34A and S31/34D prSSU-GFP mutants were transiently expressed in *Arabidopsis* protoplasts (Lamberti et al., 2011a). These experiments suggested that removal of phosphorylation sites (S31/S34A) results in plastid localization; whereas phosphomimetic S31/34D mutations lead to mislocalization of the construct (Lamberti et al., 2011a). Interestingly, western blot analysis of protoplast import suggests that the S31/34A construct shows an increase in plastid binding compared to the WT construct. Collectively our results suggest that phosphorylation is not essential for precursor import, nor does a non-dephosphorylated, phosphomimetic substrate abolish precursor import or yield mislocalization. We speculate that phosphorylation may contribute to enhanced interaction with the cytosolic guidance complex (May and Soll, 2000a) but an unidentified phosphatase removes phosphoryl groups before precursor import (Lamberti et al., 2011a). Whether this effect is simply a change in the TPs net charge or if some other more specific interaction is involved such as enhanced binding to the TOC receptors is not known. However, we have previously shown that the

phosphorylated form of a SStp-derived synthetic peptide interacts more tightly with a recombinant form of psTOC34 that would support this latter model (Schleiff et al., 2002a).

4.3.5 Potential Role of TP cis-Prolines

The TP FGLK physicochemical domains have an established role in import efficiency (Chotewutmontri et al., 2012; Lee et al., 2006; Pilon et al., 1992b); these domains often contain Pro residues (Figure 4-15). Pro residues can exist in two isomeric forms, *trans* and more uncommonly, *cis* (Song et al., 2006). The c*is* Xaa-Pro peptide bond (where Xaa can be any amino acid) can introduce a "kink" into a protein sequence, a structural element that has been suggested to contribute to several important cellular processes, including cell signaling, replication, and regulating protein activity (Andreotti, 2003; Dugave and Demange, 2003). Interestingly, TP Pro residues have been previously suggested to play a role in both recognition and efficient import of precursor proteins (Chotewutmontri et al., 2012) Furthermore, FKBP, a peptidyl proline isomerase was previously suggested as a putative guidance complex component (Fellerer et al., 2011; Lee et al., 2013). We suggest that the Pro-induced "kinks" may enable recognition or attachment of one or more proteins functioning as molecular ratchets which could facilitate efficient precursor import into the plastid stroma.

The role of the amino acid context flanking the cis-proline residues remains unclear. Do the flanking sequences simply influence the Pro isomeric form; or does recognition of the *cis* Pro isomeric form require some additional residues for its maximum specificity or recognition by the translocon(s)? Although how processing is coupled to translocation is not known, it is not surprising that this may be an intimate interaction that requires subtle spatial control between the emerging preprotein and SPP in the stroma. Investigating the potential role of *cis*-Pro racheting "kinks" in the TP sequence will be a topic of future study.

4.3.6 Conclusions

We propose a new TP import model that incorporates a dynamic recognition mechanism as depicted in Figure 4-16. TP import is likely driven by formation of a secondary structural element, which would allow TOC34 to rapidly sample the side-chain conformations of the FGLK physicochemical domain possibly regardless of orientation (Chotewutmontri et al., 2012), as the case of the reversed sequences. This dynamic

sampling would enable rapid yet relatively promiscuous binding so that many divergent TP sequences which only contain a common physiochemical property could be “trapped” via a very small class of related Toc receptors. This “trapping” is then followed by an insertion and second step of recognition by the stromal ATP-dependent molecular motors. By requiring two independent and spatially separate modes of recognition would allow the chloroplast to rapidly recognize a diverse set of proteins on the surface yet only import those that contain a recognition element for Hsp70 or an alternative ATP-dependent molecular motor. Future work is needed to see how these two domains must be spatially separated in the transit peptide to insure both rapid and precise client recognition. In addition, it is possible that these peptides may first encounter the non-polar environment of the outer membrane and form helical segments that help provide some structural “conformity” to the TPs facilitating both recognition and also lowering diffusion barriers.

CHAPTER 5: EXTENDING FUNCTIONAL ANALYSIS OF FGLK MOTIFS TO OTHER ARABIDOPSIS TRANSIT PEPTIDES

Abstract:

Even though the physicochemical FGLK motif is semi-conserved in TP sequences, and has a clear role in the import of model TPs from RuBisCO and ferredoxin, its global importance to the recognition and import of preproteins remains unclear. Our previous work has provided a solid foundation for understanding transit peptide recognition; however, due to the diversity of TP sequences, it is difficult to extend our results to represent a common mechanism for precursor recognition. We have developed a combination of bioinformatics and *in vivo* experiments to test the relative importance of FGLK domains in other *Arabidopsis* transit peptides. Our preliminary microscopy results show that this robust methodology can be used to efficiently assess TP physicochemical motifs. Our work highlights the complexity of TP structure and the interplay between different TP sequence elements.

5.1 Introduction

To investigate the universality of the FGLK motif we have developed a heuristic approach to identify other precursors containing this motif in the model plant *Arabidopsis* genome. We developed methodology using a series of online prediction tools to efficiently analyze plastid predicted proteins in the Arabidopsis genome. Our preliminary goal for this work was to identify TPs with similar length and domain structure as compared to SStp and Fdtp used in Chapter 4. This analysis can be expanded to other TPs in the future. We developed a series of mutagenesis “rules” to remove TP physicochemical domains. Next, we applied our *in vivo* protein localization assay (described in Chapter 4, see Figure 4-3) to assess the import efficiency of precursor protein constructs by using a chimeric transit peptide-fluorescent protein (FP) construct. This methodology allows us to quantify the cytosolic pool versus correctly imported protein. Using an *in vivo* approach, we can obtain high throughput, mechanistic insight about the role of the FGLK motif in the import process. Future research in the Bruce lab will focus on utilizing this toolkit of chimeric TP constructs as well as analyzing and quantifying data. Our work is expected to have a major impact in the understanding of the functional motifs in TPs.

5.2 Results

5.2.1 Genome Analysis

To start this project, our goal was to use a series of TPs with a similar length and domain structure compared to SStp and FDtp. We used a series of bioinformatics tools to narrow our *Arabidopsis* protein dataset to a small set of proteins for our preliminary studies (Figure 5-1). Analysis of the *Arabidopsis* genome found 912 highly confidently predicted (using multiple prediction algorithms) plastid precursors, of which, 327 proteins contain a similar length of transit peptide at ~50-60 a.a. and 231 precursors contain at least one 'FGLK' motif. We selected a subset of 24 proteins based on the presence of FGLK motif around residues 28-39. Using multiple online tools these precursors were then further analyzed to predict their subcellular location. The output from these prediction tools was integrated into a mathematical scoring algorithm (described below) to select the most confidently predicted chloroplast proteins, while avoiding potentially dual-targeted proteins to the mitochondria, shown in Figure 5-2. Figure 5-1 outlines our methodology for selecting TPs for this study. To limit the number of inconsistencies found in the localization prediction software, we used an algorithm to weigh the localization software predictions to avoid the mitochondria, while increasing the chances of importing to the chloroplast. This involved consolidation of the following prediction tools: ChloroP 1.1, TargetP, iPSORT, Predotar, PredSL, ProtComp, and Protein Prowler (Figure 5-2). The number generated from this algorithm allowed us to identify 7 proteins that were most confidently predicted to the chloroplast: Glutamine-5-phosphoribosylpyrophosphate amidotransferase, Glucose-6-phosphate dehydrogenase, CHLI subunitol magnesium chelatase, 2C-methyl D erythritol 2,4- cyclophosphate synthase, Beta-carotenoid hydroxylase, Allene-oxide cyclase, and Rubisco Activase (Figure 5-2). Figure 5-3 shows the amino acid sequence of each TP used in this study. Proteins were ranked 1-7 based on their algorithm score, which corresponds to their likelihood to be targeted to the chloroplast.

Figure 5-1: Selection of additional TP precursors

Analysis of the *Arabidopsis* genome found 912 highly confident predicted plastid precursors, of which, 327 proteins contain a similar length of transit peptide at ~50-60 a.a. and 231 precursors contain at least one 'FGLK' motif. We selected a subset of 24 proteins based on the presence of FGLK motif around residues 28-39, the functional annotation (Plant Proteome Database and Phytozome), proteomic localization (Plant Proteome Database), level of expression (*Arabidopsis* eFP Browser), and coexpression analysis (ATTED-II). These precursors were then further analyzed using localization predictions (TargetP, ChloroP 1.1, iPSORT, Pedotar, PredSL, ProtComp, Protein Prowler). The output from these prediction tools was integrated into a mathematical scoring algorithm to select the 7 most confidently predicted chloroplast proteins, while avoiding potentially dual-targeted proteins to the mitochondria.

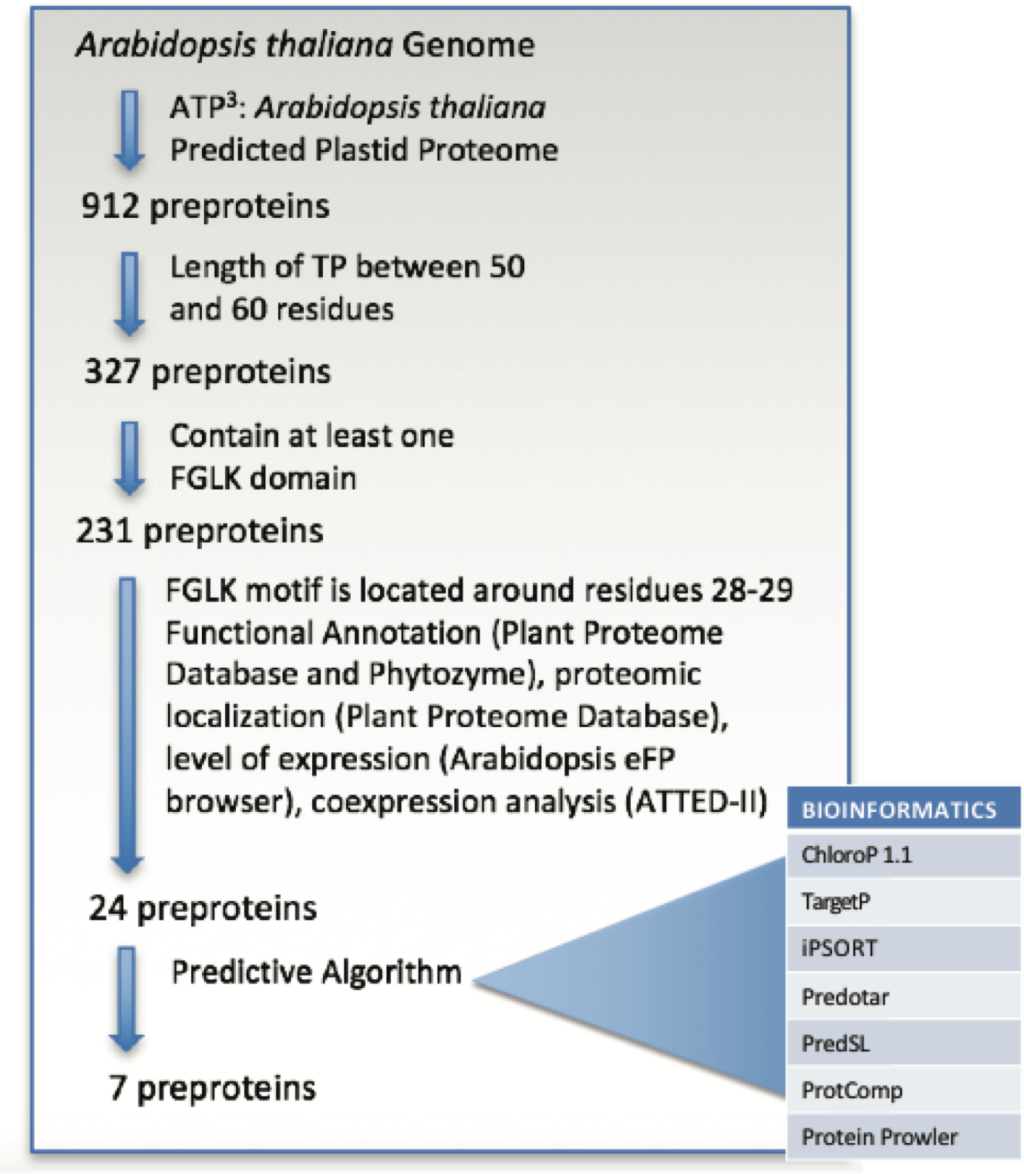


Figure 5-1 Continued

Figure 5-2: Predicting proteins using online plastid prediction algorithms

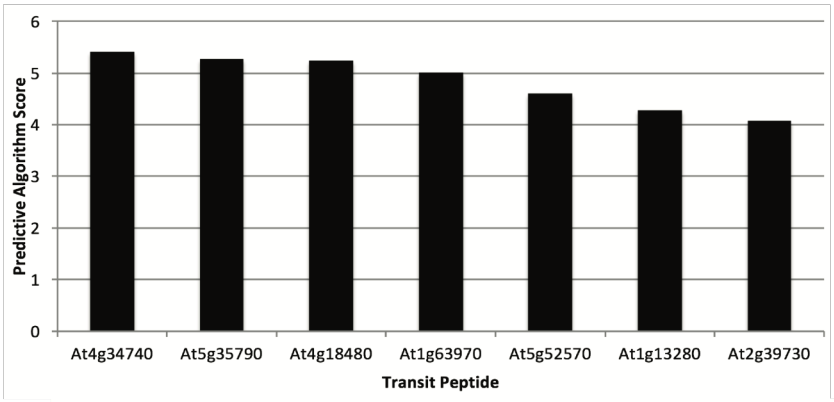
To limit the number of inconsistencies found in the localization prediction software, an algorithm weighting the localization software was developed to avoid the mitochondria, while increasing the chances of importing to the chloroplast. This involved consolidation of the following prediction tools: ChloroP 1.1, TargetP, iPSORT, Predotar, PredSL, ProtComp, and Protein Prowler. The number generated from this algorithm allowed us to choose 7 proteins that were most confidently predicted to the chloroplast: Glutamine-5-phosphoribosylpyrophosphate amidotransferase, Glucose-6-phosphate dehydrogenase, CHLI subunitol magnesium chelatase, 2C-methyl D erythritol 2,4- cyclophosphate synthase, Beta-carotenoid hydroxylase, Allene-oxide cyclase, and Rubisco Activase.

A

ChloroP 1.1	Chloroplast Localization Cleavage Site (CS-Score) "cTP" "cTP-length"	[Chloro loc]
TargetP	Length cTP mTP SP Other RC	[cTP – (mTP + SP + other)] – TargetP [RC]
iPSORT	Localization Target Sequence	[-1 for mito or +1 for chloro]
Predotar	Chloroplast Localization Mitochondria Localization ER Localization Other Localization	[chloro - (mito+ ER+ other)
PredSL	Chloroplast Localization Mitochondria Localization Secreted Protein Cleavage Site	[chloro loc- (mito loc + SP loc)]
ProtComp	Chloroplast Localization Mitochondria Localization Plasma Membrane Nuclear Localization Peroxisome Localization	[chloro loc- (mito loc + PM loc + nuc loc+ perox loc]/9
Protein Prowler	Chloroplast Localization Mitochondria Localization Secreted Protein Other Localization	[chloro loc- (mito loc + SP loc + other loc)]

Predictive Algorithm = [chloro loc] + [cTP – (mTP + SP + other)] – [RC] + [-1 for mito or +1 for chloro] + [chloro - (mito+ ER+ other)] + [chloro loc- (Mito loc + SP loc)] + [chloro loc- (Mito loc + PM loc + nuc loc+ perox loc]/9 + [chloro loc- (mito loc + SP loc + other loc)]

B



C

Gene ID	Gene Name	Score
At4g34740	Glutamine-5-phosphoribosylpyro phosphate amidotransferase	5.408
At5g35790	Glucose-6-phosphate dehydrogenase	5.279
At4g18480	CHLI subunitol magnesium chelatase	5.238
At1g63970	2C-methyl D erythritol 2,4-cyclodiphosphate synthase	5.007
At5g52570	Beta-carotenoid hydroxylase	4.611
At1g13280	Allene-oxide cyclase	4.274
At2g39730	Rubisco activase	4.073

Figure 5-2 Continued

5.2.2 Mutagenesis and cloning strategy

To test our representative array of TPs, we designed a series of mutagenesis rules to assess the contribution of the physicochemical domains of the TPs. The rules are shown in Figure 5-4. The mutagenesis was designed to remove either one (Rule 2 & Rule 3) or both (Rule 4) FGLK domains. We also added a rule to test the global importance of basic residues in the FGLK domain (Rule 5; Figure 4-6, Figure 5-4). Our miscellaneous rule was added to test any additional physicochemical features we wanted to test in our analysis. The selected mutants for each TP are shown in Figure 5-5. Regions selected for mutagenesis are highlighted in yellow. We mutated regions of interest to Ser because Ser is the most common amino acid in TP sequences (Chotewutmontri et al., 2012; Holbrook et al., 2016), so we expect this mutation would have minimal effect on TP secondary structure. We ordered each TP "Rule Set" as a single, synthetic, concatemerized sequence (6 vectors in total) (Epoch Biotech, Missouri City, Texas) and designed primers so that each TP could be moved into an expression construct in frame with YFP as a reporter construct. Our cloning strategy is described in Figure 5-6. Briefly, in order to move each TP sequence from its concatemerized rule vector, we designed primers corresponding to each TP within the vector. Each TP was moved into an individual CloneJet vector (42 unique CloneJets) then cloned into an expression vector in frame with YFP. As described in detail in Chapter 4, our expression vector setup allows us to use biolistic transformation of onion epidermal cells to test the localization of TP-YFP mutants. Our CloneJet intermediate vector system will allow us to easily move our constructs into binary vectors for expression using *Agrobacterium tumefaciens* transformation.

5.2.3 *In vivo* analysis of TP6: Allene oxide cyclase

We have started preliminary testing of TP6 in our biolistic assays. This set of TP-YFP mutants correspond to the precursor of Allene-oxide cyclase (AOC) and is implicated in the biosynthesis of jasomonate. Qualitative results are shown in Figure 5-7. We plan to test each mutant using both qualitative and quantitative microscopy analysis. So far, we have successfully cloned 5 of the 6 TP6 constructs. Interestingly, Rule 6 (Miscellaneous) shows a very pronounced abrogation of import as compared to any of the other mutants. We have highlighted the unique mutagenized sequence in Rule 6 (Figure 5-7, bottom

panel). Although we cannot make conclusions about the import competence of TP6 until we have the full selection of mutants, it is clear that the mutants in this assay shows a range of import capabilities.

5.3 Discussion

5.3.1 Development of bioinformatics and microscopy methodolgy

Using a heuristic bioinformatics' approach we have selected 7 new chloroplast transit peptides to perform a combined directed mutagenesis and *in vivo* targeting assay. These peptides share a similar length and targeting prediction to the well-studied model transit peptides, SStp and Fdtp. These sequences also contain two of the semi-conserved FGLK motifs. Furthermore, these sequences correspond to proteins with a wide range of function; from hormone synthesis to components involved in photosynthesis. Classic model TPs used in transport research have been largely limited to highly imported photosynthetic proteins. Our goal is to establish that the FGLK physicochemical domain has a broad importance in plastid protein import regardless of protein function.

We generated a set of synthetic genes to make six distinct classes of mutations. These synthetic TPs have now been cloned in frame with the fluorescent protein YFP. We have constructed the transient expression vectors to allow biolistic transformation into onion epidermis tissue, which contain non-green leucoplasts. This system has the advantage of being free from chlorophyll auto fluorescence and permits a more direct quantitation of the cytosolic/plastid localization. However, in the future we will use our new Leica SP8 LSCM that is equipped with both a white light laser and the new LightGate fluorsecence lifetime discrimination feature. This combination will allow this assay to be applied to pea seedlings that contain differentiated chloroplasts. This will involve an *Agrobacterium*- mediated transformation (discussed in Chapter 6) that will allow large quantities of isolated chloroplasts to be isolated and biochemically/immunologically characterized for both import and processing fidelity. Furthermore, future analyses will permit a detailed comparison of the targeting activity and fidelity on two distinct classes of plastids that may in fact have distinct TOC translocon composition with different TOC34/TOC159 isoforms.

Glutamine-5-phosphoribosylpyro phosphate amidotransferase	MAATSSISSSLSLNAKPNKLSNNNNNNKPHRFLRNPFLNPSSSSFSPLPASISSSSSPSFPL
Glucose-6-phosphate dehydrogenase	MATHSMIIPSPSSSSSSLATAASPFKETLPLFSRSLTFPRKSLFSQVRLRFFAEKHSQLD
CHLI subunitol magnesium chelatase	MASLLGTSSSAIWASPSLSSPSSKPSSSPICFRPGKLFGSKLNAGIQIRPKKNRSRYHVSVMNVATEINS
2C-methyl D erythritol 2,4-cyclodiphosphate synthase	MATSSTQLLLSSSSLFHSQITKKPFLLPATKIGVWRPKKSLSLSCRPSASVSAASSAVDVNE
Beta-carotenoid hydroxylase	MAAGLSTIAVTLKPLNRSSFSANHPISTAVFPPSLRFNGFRRRKILTVCFVVEERKQSSPMD
Allene-oxide Cyclase	MIMASSAAASISMITLRNLSRNHQSHQSTFLGFSRSFHNQRISSNSPGLSTRARSTTSSTGG
Rubisco activase	MAAAVSTVGAINRAPLSLNGSGSGAVSAPASTFLGKKVVTVSRFAQSNKKSNGSFKVLAVKEDKQTDG

Figure 5-3: Preprotein amino acid sequences

The amino acid sequence of selected preproteins. We used a prediction parameter from Chloro P to determine the proposed length of each transit peptide (c-TP= 50- 60). The signaling sequence for each preprotein is shown in black text. The red text indicates the first 10 residues of the mature domain, following each predicted transit peptide region.

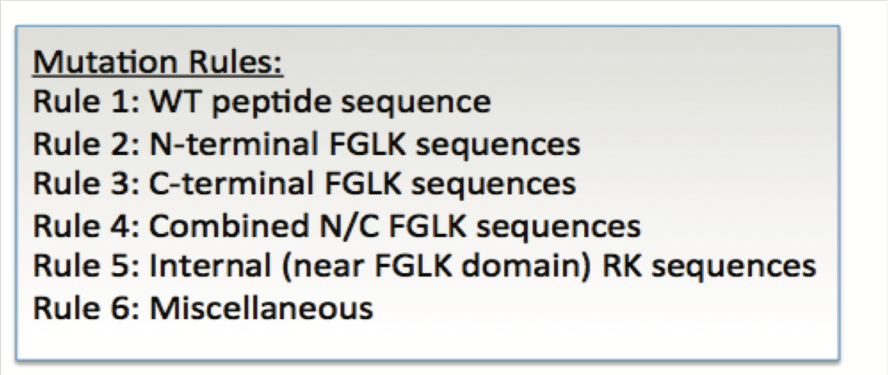


Figure 5-4: Mutagenesis Rules

These rules indicate the physicochemical regions selected for systematic serine mutagenesis chosen for each preprotein. These were determined based on the results from our algorithm predictions, as well as a sequence alignment of each TP. We utilized our previous work (Chapter 4) as a basis to determine which physicochemical domains to focus on. Miscellaneous Rule 6 was included to cover any unique physicochemical domains we wanted to include in our mutagenesis studies.

Figure 5-5: Selected TP Mutants

Each panel shows the amino acid sequence for a selected TP sequence for precursors 1-7.. The yellow regions indicate regions that were mutated from their wild type residue to a Serine residue.

Figure 5-5 Continued

5.3.2 Identification of FGLK sequences and complexity of TP structure

We used bioinformatic programs to identify TP FGLK sequences in our selected TPs. Importantly, this analysis highlighted the complexity of TP physicochemical domains. We attempted to limit our analysis to TPs with similar properties (2 FGLK domains, close to equal length of TP). However, due to the loose conservation of the FGLK domain, some TPs had additional sequences that may act as an FGLK sequence. For example, we find that TP6-Rule6 (Figure 5-7) shows strongly abrogated import when the sequence "RNLSRNHQSHQ" is removed. It is possible that this sequence is important due to its positive charge; RNHQ may also act as a a section of a surrounding FGLK physicochemical domains. Thus, although it is clear that FGLK has a central role in TP recognition and import, the significance of TP structure still requires further study. This assay highlights the complexity of TP structure and the interplay between different sequence elements. Our future analyses will assay TP1-TP6 for their import capability. The robust methodology of this study allows us to easily expand to study additional TPs with different domain structure, or to expand our mutagenesis rule library for TP1-TP6. We can easily extend our analysis to analyze unique physicochemical motifs identified in other assays.

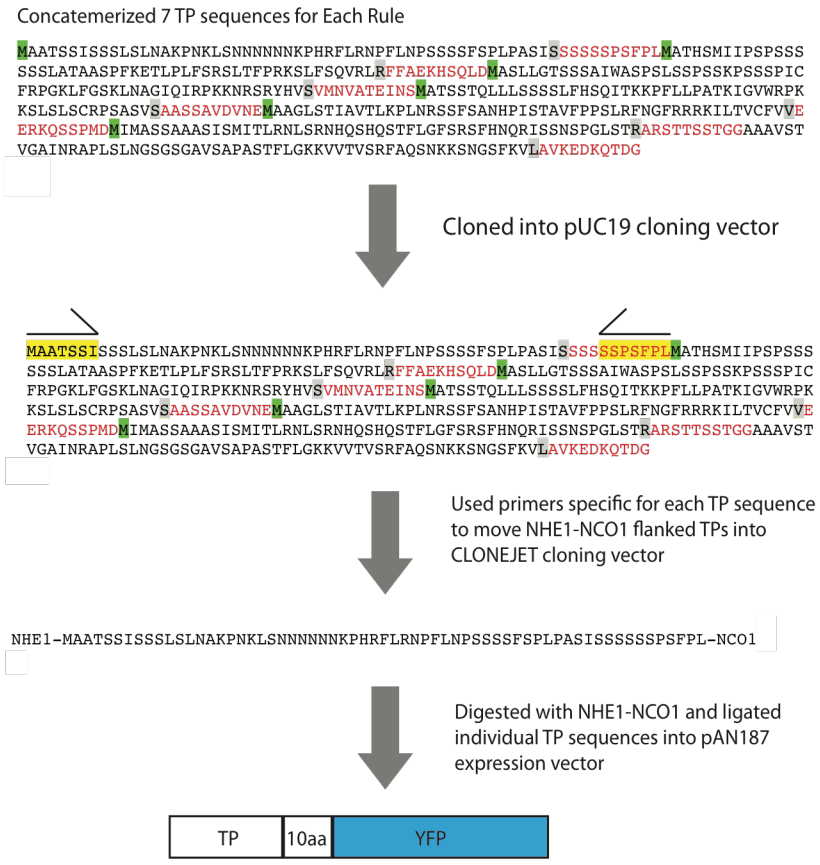
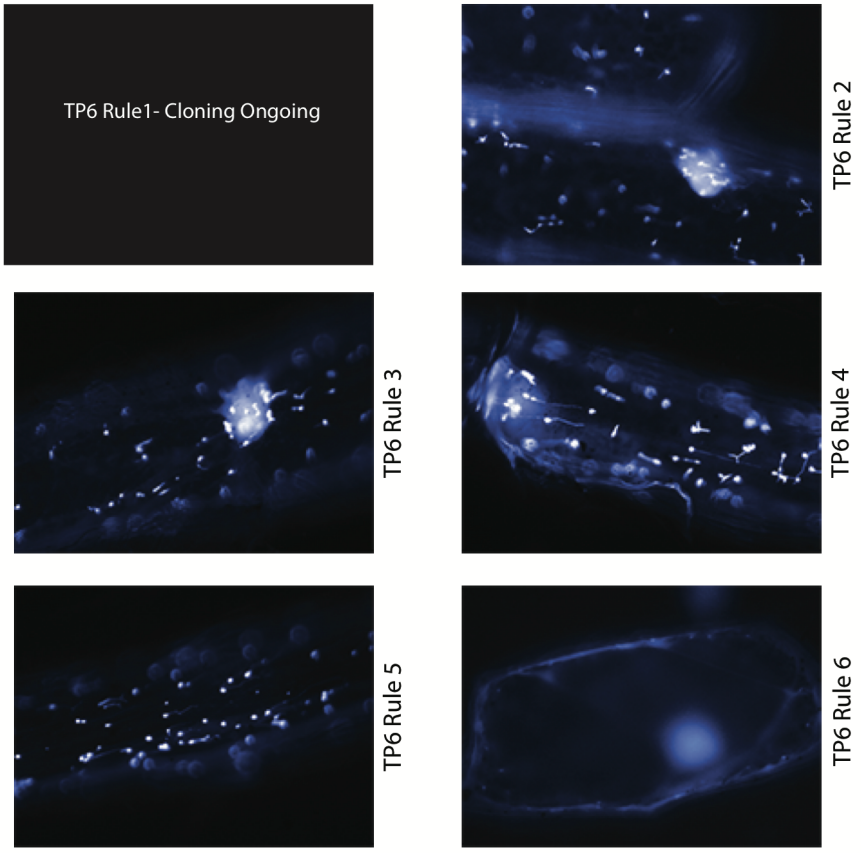


Figure 5-6: TP molecular cloning and expression strategy

Each transit peptide was ordered in a concatemerized vector per rule (Epoch Biotech). In order to move each TP individually from its Rule vector, we designed primers corresponding to each TP within a rule vector. Each rule was stored in a CloneJet vector. Then, TPs were incubated with restriction enzymes prior to their placement in an expression vector to ensure unidirectional integration into the vector. Each final construct consists of an individual TP sequence, 10 amino acids of the precursor mature domain, and in-frame YFP as a fluorescent reporter.



Allene Oxide Cyclase (At1g13280) (EI= 4.274; c-TP= 52)

```
Rule 1: MIMASSAAASISMITLRNLSRNHQSHQSTFLGFSRSFHNQRISSNSPGLSTRARSTTSSTGG
Rule 2: MIMASSAAASISMITLRNLSRNHQSHQSTSSSSSSFHNQRISSNSPGLSTRARSTTSSTGG
Rule 3: MIMASSAAASISMITLRNLSRNHQSHQSTFLGFSRSSSSSISSNSPGLSTRARSTTSSTGG
Rule 4: MIMASSAAASISMITLRNLSRNHQSHQSTSSSSSSSSSSSISSNSPGLSTRARSTTSSTGG
Rule 5: MIMASSAAASISMITLRNLSRNHQSHQSTFLGFSSSFHSSSISSNSPGLSTRARSTTSSTGG
Rule 6: MIMASSAAASISMITLSSLSSSSSSSSTFLGFSSSFHSSSISSNSPGLSTSARSTTSSTGG
```

Figure 5-7: *In vivo* qualitative analysis of TP6

A. Representative epi-fluorescence microscopy images of TP6 Rule2-Rule6. Transient expression of chimeric TP6-YFP mutant constructs in onion (*Allium cepa*) were live-imaged using epi-fluoresecence microscopy. Bottom panel depicts alignment of TP6 sequence with each mutant region shown in red. The unique mutagenized sequence of TP6-Rule6 is highlighted in yellow.

CHAPTER 6: IMPROVED APPROACHES FOR VISUALIZATION, QUANTIFICATION, AND DETECTION OF *IN VIVO* PLASTID PROTEIN TARGETING

Abstract:

Correct protein localization is integral to the function of biological systems. Transient fluorescent protein expression represents a simple, high-throughput method to assay protein localization in living cells. We have tested and improved several popular transient expression techniques using the primary reference organism for *in organello* localization studies, dwarf pea seedlings (*Pisum sativum*). For our comparative analysis, we have developed new methodology for *Agrobacterium tumefaciens* based transformation of dwarf pea seedlings (*Pisum sativum*). We are also developing protocols for transient protein expression in pea protoplasts. We are testing many different parameters, including plant age, culturing technques and transformation protocols.

We extended methodology development to improve biolistic protocols for monocot plants. We have focused on developing protocols using commercial leeks (*Allium ampeloprasum var. porrum*) to test plastid localization of fluorescent proteins. We find that use of commercial leek epidermis tissue significantly improves biolistic transformation efficiency compared to commercial onion epidermis tissue (*Allium cepa*). Although the physiology of the leek leads to a developmental gradient of leucoplasts to chloroplasts, leek tissues can be transformed at a consistent rate. Furthermore, laser scanning confocal microscopy (LSCM) analysis using LightGate allows us to move our analysis into the most physiologically relevant plastid import system, differentiated chloroplasts. We discuss our results in context of classical methodologies compared to these technical advances.

6.1 Introduction

Although protein trafficking has been the focus of research for many years, our understanding of the factors that influence protein localization is far from complete. Transient expression techniques allow for relatively high-throughput analysis of fluorescent protein localization. Our research is focused on protein transport into the plant cell plastid. Plastids represent a particularly complex destination, requiring several translocons located in three membranes of this organelle. To assess the localization of WT/mutant plastid proteins *in vivo,* we have developed an assay using a chimeric transit peptide (TP)-fluorescent protein (FP) construct to analyze their cellular sublocalization using fluorescence microscopy. We have already developed methodology that allows us

to quantify the cytosolic pool versus correctly imported protein (discussed in Chapter 4). We recently received access to the Leica SP8 LightGate microscope, with laser scanning confocal microscopy (LSCM) technology that greatly improves our ability to do complicated imaging. Previously, we were technically limited to analysis of protein import in non-green plastids (leukoplasts) due to the color battle between high chlorophyll autofluorescence in green differentiated chloroplasts versus the fluorescence of our chimeric TP reporter construct. Fluorescence is defined by both emission intensity and fluorescence lifetime. The combination of a white light laser and new technology called LightGate allows us to filter out reflected light based on fluorescence lifetime, and thus detected signal is purely fluorescent protein emission. Thus, this technology allows us to move our research into the most physiologically relevant system, differentiated chloroplasts.

In our previous research (Chotewutmontri and Bruce, 2015; Chotewutmontri et al., 2012; Holbrook et al., 2016), we chose to use onion epidermal cells because their lack of chlorophyll was amendable to our quantification method. However, the most robust way to perform both *in organello* and *in vivo* experiments would be to use the same system (differentiated chloroplasts derived from dwarf pea seedlings) for both methodologies – however, protocols are not currently available. Thus, we focused on developing and improving methodologies used for plastid research. Several techniques are commonly used for transient plant expression: *Agrobacterium tumefaciens* mediated transformation, protoplast transformation and biolistic transformation. We focused on comparing the protocol ease, transformation rate, and quantification efficacy to compare methodologies. Our TP-FP construct combines the wild type transit peptide of ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCo) of tobacco with 20 amino acids of the mature domain to ensure WT protein behavior, with C-terminal YFP as a fluorescept reporter. Because the behavior of this construct has been analyzed in tobacco, *Arabidopsis thaliana* and onion plastids, we used this YFP chimera as an output to directly compare the methodologies tested in this study.

6.2 Results

6.2.1 Plant growth

Leeks and onions used in our biolistic analysis were purchased commercially on the day of the experiment. Using sterile conditions, onion inner epidermal peels from the third bulb layer (about 2.5 cm x 2.5 cm) were placed on agar plates containing 1x Murashige and Skoog (MS) salts, 30g/L sucrose and 2% agar. Leek sections were removed from the bulb with epidermis intact, and placed on MS agar plates. Adaxial onion epidermal peels or leek sections were bombarded within 1 hour of placing on agar plates. Biolistic bombardment is described later in this Chapter.

Dwarf pea seeds were acquired from J.W. Jung Seed Company and stored at 4C. Seeds were imbibed overnight by running cold tap water with aeration at room temperature before being planted. Seeds were placed on top of "coarse" grade vermiculite in a plastic pot. The seeds were then covered with a 1.5cm layer of vermiculite. A wire mesh screen with 0.6 mm x 0.6 mm opening was placed on top of the seeds and secured by folding the mesh around the edges of the pot. Peas were grown in a growth chamber with illumination at approximately 100 uE/m^2/sec on a 14h light and 10 h dark cycle. The temperature was set at 18C and 20C for the light and dark periods, respectively. We tested both 9 day old seedlings and 11 day old seedlings and found that 11 day old seedlings transformed the most effectively using *Agrobacterium tumefaciens*. Our 11 day old seedlings had about 2 sets of compound leaves, with both sets of leaves expanding. Protoplasts were made using Day 14 leaves. Figure 6-1 shows an example of the setup used for pea growth on Day1, Day6 and Day 14.

6.2.2 Productions of Protoplasts

On Day 14, 1.5g of pea leaves were collected and the epidermal layer was carefully removed using forceps. Pea tissue was cut into approximately 5mm strips and were placed into 25mL Enzyme Solution (1.5% Cellulase -*Trichoderma viride*, 0.75% Macerozyme R-10, 0.6M Mannitol, 10mM MES at pH 5.7, 10mM CaCl2, and 0.1% BSA). The pea tissue and enzyme mixture was gently rotated for 4 hours in light at room

Figure 6-1: *Pisum sativum* growth

Dwarf pea seedlings were acquired from J.W. Jung Seed Company and stored at 4C. Seeds were imbibed overnight by running cold tap water with aeration at room temperature before being planted. Seeds were placed on top of “coarse” grade vermiculite in a plastic pot (A). The seeds were then covered with a 1.5cm layer of vermiculite. A wire mesh screen with 0.6 mm x 0.6 mm opening was placed on top of the seeds and secured by folding the mesh around the edges of the pot. Peas were grown in a growth chamber with illumination at approximately 100 uE/m^2/sec on a 14h light and 10 h dark cycle. The temperature was set at 18C and 20C for the light and dark periods, respectively.

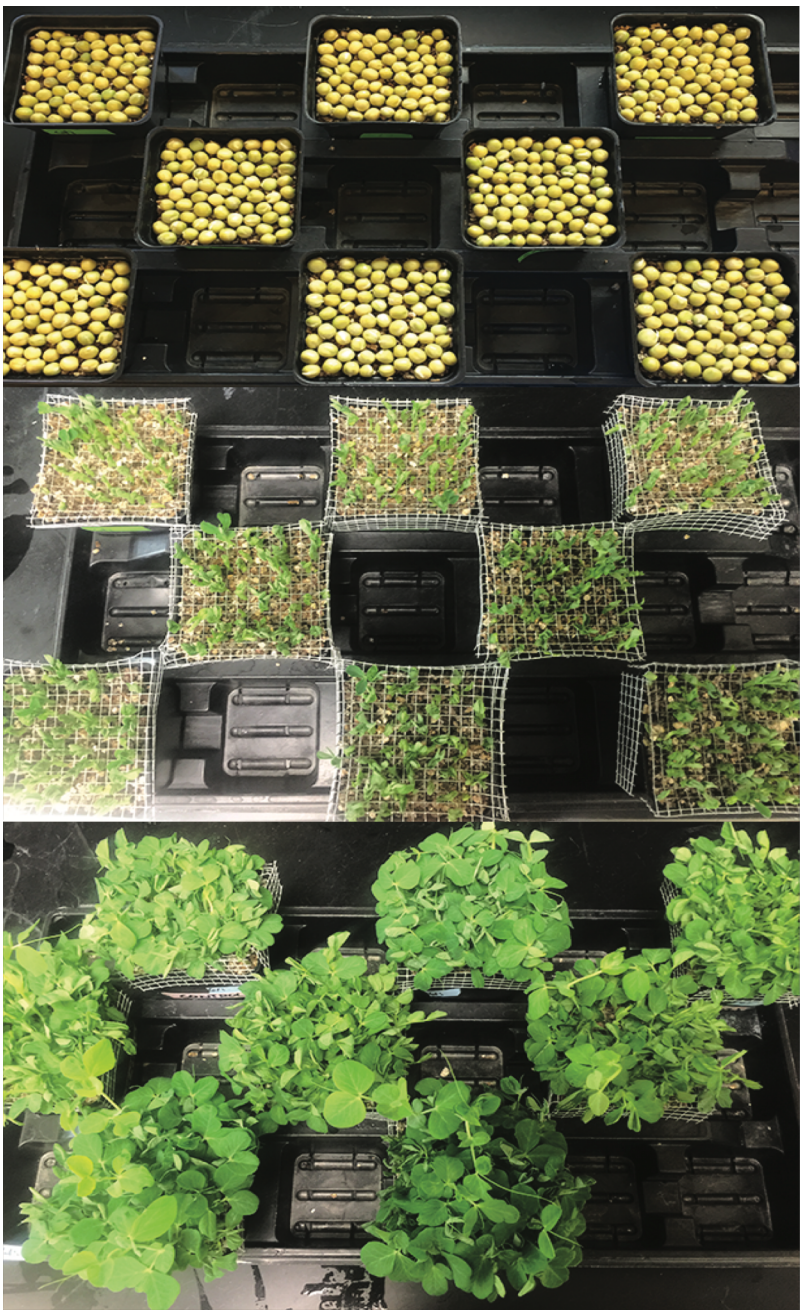


Figure 6-1 Continued

temperature. Once the enzymatic digestion was complete, equal parts of W5 Solution (154 mM NaCl, 125 mM CaCl2, 5 mM KCl and 2 mM MES at pH 5.7) was added to the flask and was shaken vigorously for 15 seconds to neutralize the enzymes. The mixture was filtered through 40µm nylon mesh, and the mesh was thorough washed with fresh W5 Solution. Filtrate containing protoplasts was centrifuged in a swinging bucket rotor at 1,500 rpm for 3 minutes. The supernatant was discarded, and the pellets were suspended in 30ml of W5 solution. Resuspended protoplasts were centrifuged again, and the pellet was gently resuspended in MMG solution (0.4 M mannitol, 15 mM MgCl2 and 4 mM MES at pH 5.7). Using a hemocytometer, we counted protoplasts; the final protoplast concentration can be adjusted accordingly for downstream experiments. A panel of example images of pea protoplasts are shown in Figure 6-2.

6.2.3 *Agrobacterium tumefaciens* mediated transformation of *Pisum sativum*

Agrobacterium tumefaciens based transformation is a useful system for transient expression in systems like *Arabidopsis* and tobacco, but protocols to transform *Pisum sativum* are not well developed. We used Arabidopsis floral dip media for our study (0.22% (w/v) Murashige & Skoog Basal Medium, 5.0% (w/v) sucrose, 44 nM benzylaminopurine, pH 5.7 (Clough and Bent, 1998). We utilized strain GV3101 *Agrobacterium tumefaciens* (kindly provided by Dr. Andreas Nebenfuehr, University of Tennessee) carrying a binary plasmid encoding our TP-FP construct. Agrobacterium was cultured in YEB media (10g/L yeast extract, 5g/L beef extract, 5g/L peptone, 5g/L sucrose, 0.5g/L $MgSO_4$ $7H_2O$) containing the correct antibiotic at 28C and shaking at 225 rpm for 20-24h until stationary phase. Cells were harvested by centrifugation at 3,500 g for 10 minutes and washed once with ice cold water and once with infiltration media. Cells were resuspended to OD_{600} of 0.1 for infiltration. Surfactant Silwet L-77 (Lehle Seeds, Round Rock, TX, USA) at 0.05% was added to the cells immediately prior to vacuum application. To transform pea seedlings, the top vermiculite layer of the pea seedling pot was shaken off completely. Vermiculite causes bubbles during the vacuum infiltration step and affects transformation efficiency. The pea pot was then inverted and the tissues were submerged in the *Agrobacterium* suspension and placed in the vacuum chamber. The vacuum (approximately 9.15 kPa) was held for 90 s and released; this process was repeated once. The pot was then wrapped in a plastic bag to maintain humidity. The

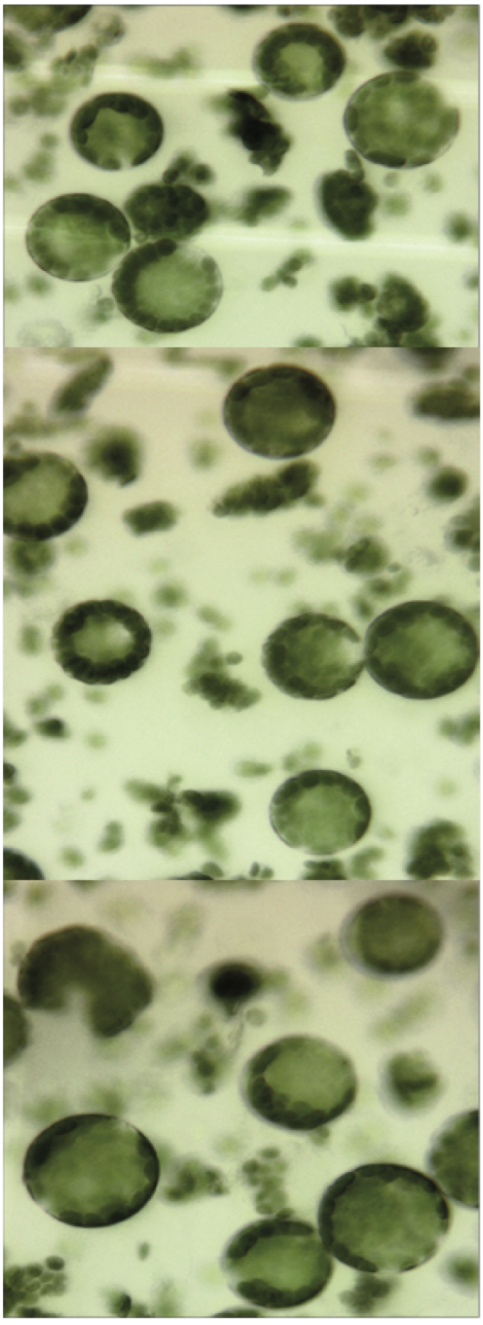


Figure 6-2: Protoplasts derived from *Pisum sativum*

Panel shows protopasts produced from 14 day old pea seedlings imaged with light microscopy.

plants were returned to the growth chamber and plastic bags were removed after 24h. This transformation sequence is shown in Figure 6-3.Leaves were analyzed for fluorescence after 48 hours using a Leica SP8 confocal microscope; our analysis suggests that optimal visualization occurs between 48-72 hrs (data not shown). We visualized at pea leaves at 48 hrs post transformation using confocal microscopy (Figure 6-4).

6.2.4 Biolistic transformations

All assays were performed essentially as described in (Chotewutmontri et al., 2012) and (Nelson, 2007). The PDS-1000/He Biolistic Particle Delivery System (Bio-Rad) was used. 60 mg/mL of M-17 tungsten particles (onion) in 50% glycerol were prepared by vortexing with 70% ethanol for 10 min and rinsing 3 times with sterile water. The plasmid vectors (pAN187 vectors) were coated on tungsten particles by incubating 10 μL of the tungsten solution with ~1 μg of plasmid DNA (<10 μL), 25 μL of 2.5 M $MgCl_2$ and 5 μL of 200 mM spermidine. The coated particles were incubated for 20 min on ice, washed once in 70% ethanol and 3 times in absolute ethanol before being resuspended in 10 μL of absolute ethanol. Macrocarriers (Bio-Rad) were coated with 10 μL of the tungsten-DNA mixture. Bombardment was performed using 1,100 psi overspeed discs (Bio-Rad). As described in Section 6.2.1, adaxial epidermal peels or leek sections with intact epidermal sections were used where the adaxial surface was placed against the surface of a Murashige and Skoog Basal Media agar plate and kept at room temperature in the dark until analysis (14-16 hrs post bombardment unless otherwise noted).

6.2.4 Leek epidermis transforms with a higher efficiency than onion epidermis

We have focused on improving existing biolistic transformation protocols used in the Bruce lab. Initially, we used previously established biolistics protocols to transform onion epidermis (Nelson, 2007). After 12-18 hours, we imaged this cells using epifluorescence microscopy and measured the distribution of YFP to generate a ratio of plastid:cytosol YFP signal localization (Chotewutmontri et al., 2012; Holbrook et al., 2016). This methodology allows us to indirectly determine levels of protein import to the plastid compared to the surrounding cytosol (Figure 4-3). A fluorescence ratio helps control for cell-to-cell variations in expression levels following biolistic transformation. When mutant

constructs are not correctly targeted to the plastid, we observe YFP fluorescence in the cytosol and other cellular compartments including the nucleus. Our current study improves several aspects of transient protein expression using commercial leeks (*Allium ampeloprasum*).

Instead of forming a tight bulb like an onion, leeks produce a long cylinder of bundled leaf sheaths; there is a clear developmental greening along the leek cylinder. This can be advantageous, since both undifferentiated leukoplasts as well as differentiated chloroplasts are available for transformation. To test biolistic transformation, on leek epidermis tissue, we divided commercial leeks into three sections: Bottom (B), containing undifferentiated leucoplasts; Middle (M), containing a mixture of leucoplasts and chloroplasts; and Top (T) containing mostly differentiated chloroplasts (Figure 6-5A). We first tested the transformation rate of leek epidermal tissue from Section M compared to onion epidermal tissue. We have previously characterized the highly reproducible plastid localization of this construct. We analyzed 5 sections (about 1.5×10^5 microns^2 each) and find that leeks show a significantly higher rate of transformation ($p<0.0001$) as compared to onion epidermis (Figure 6-5B-C and quantification shown in 6-5D). We tested several aspects of transient protein expression. First we assessed the chlorophyll content of 5 leek sections (Figure 6-5A; 6-6A). Sections 4 and 5 correspond to “Top” region of the leek; section 3 corresponds to the “Middle” section, and sections 1 and 2 correspond to the “bottom” section. Differentiated chloroplasts contain higher amounts of chlorophyll; there is a clear gradient in plastid differentiation based on the content of chlorophyll. We extended this analysis by assessing the protein content using BCA analysis (Thermo Fisher). We used RIPA buffer solubilization by dry weight of the section and graphed the measured total protein concentration. We find that the sections have a nearly equal protein content (Figure 6-6B). After initially observing that the rate of transformation was greatly improved in commercial leeks versus onions, we decided to more carefully test plastid protein localization. We set up a series of basic fluorescence based localization assessments. First, we tested whether there was difference in plastid protein localization based on the section location within the bulb, as depicted in Figure 6-7A. We transformed the sections using biolistic transformation and quantified protein localization of WT SStp-YFP.

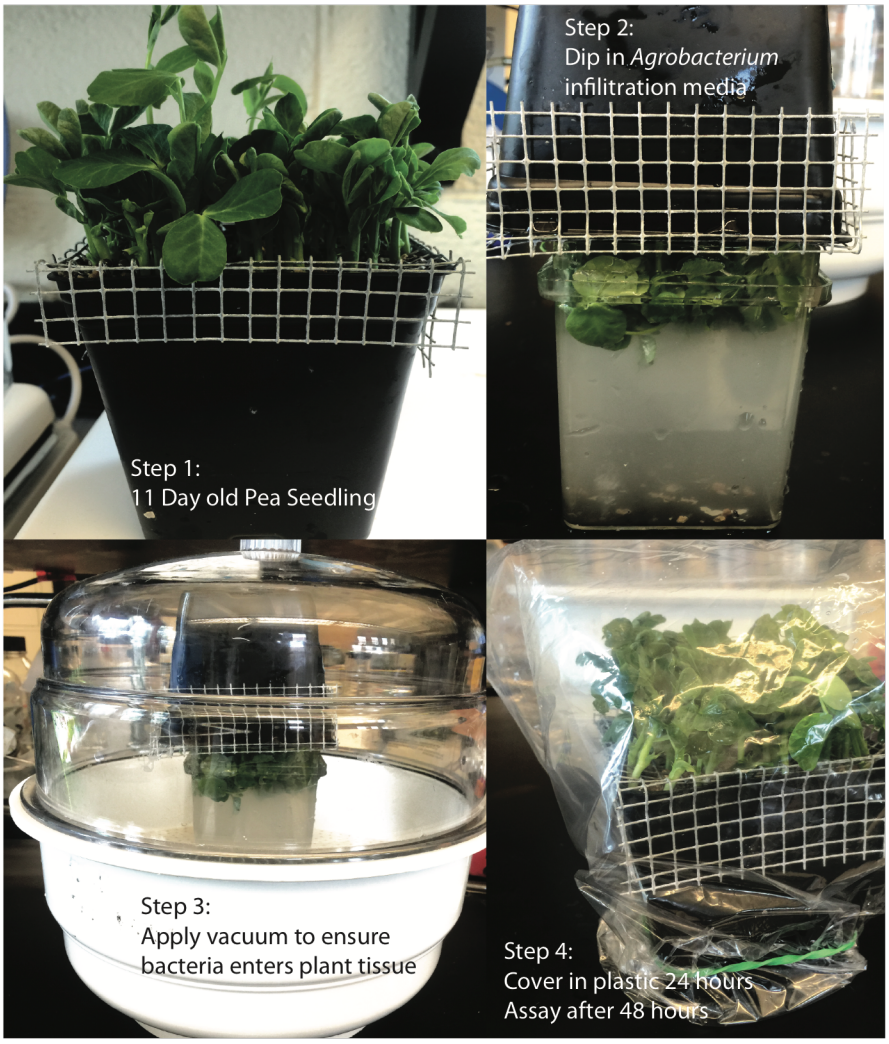


Figure 6-3: Transformation of pea tissue using *Agrobacterium tumefaciens*

11 day old pea seedlings (1) were dipped into infilitration media containing 0.1 OD_{600} of Agrobacterium tummefaciens harboring our gene of interest (2). To ensure leaf infiltration, the dipped plant was placed into a chamber and vacuum was applied twice for 90s (3). The vacuum infiltrated plant was covered in plastic for 24 hours to maintain humidity.

Figure 6-4: Analysis of chloroplasts transformed with SStp-YFP using LSCM

Agrobacterium transformed pea leaves expressing WT SStp-YFP and imaged using LightGate technology on the Leica SP8 confocal microscope. Chlorophyll fluorescence is shown in red and YFP fluorescence is shown in yellow. Bottom right panel shows a schematic of a leaf crossection. White and yellow arrows correspond to transformed cells. Next we will develop methodology to accurately quantify the subcellular localization of TP-YFP proteins expressed in differentiated chloroplasts.

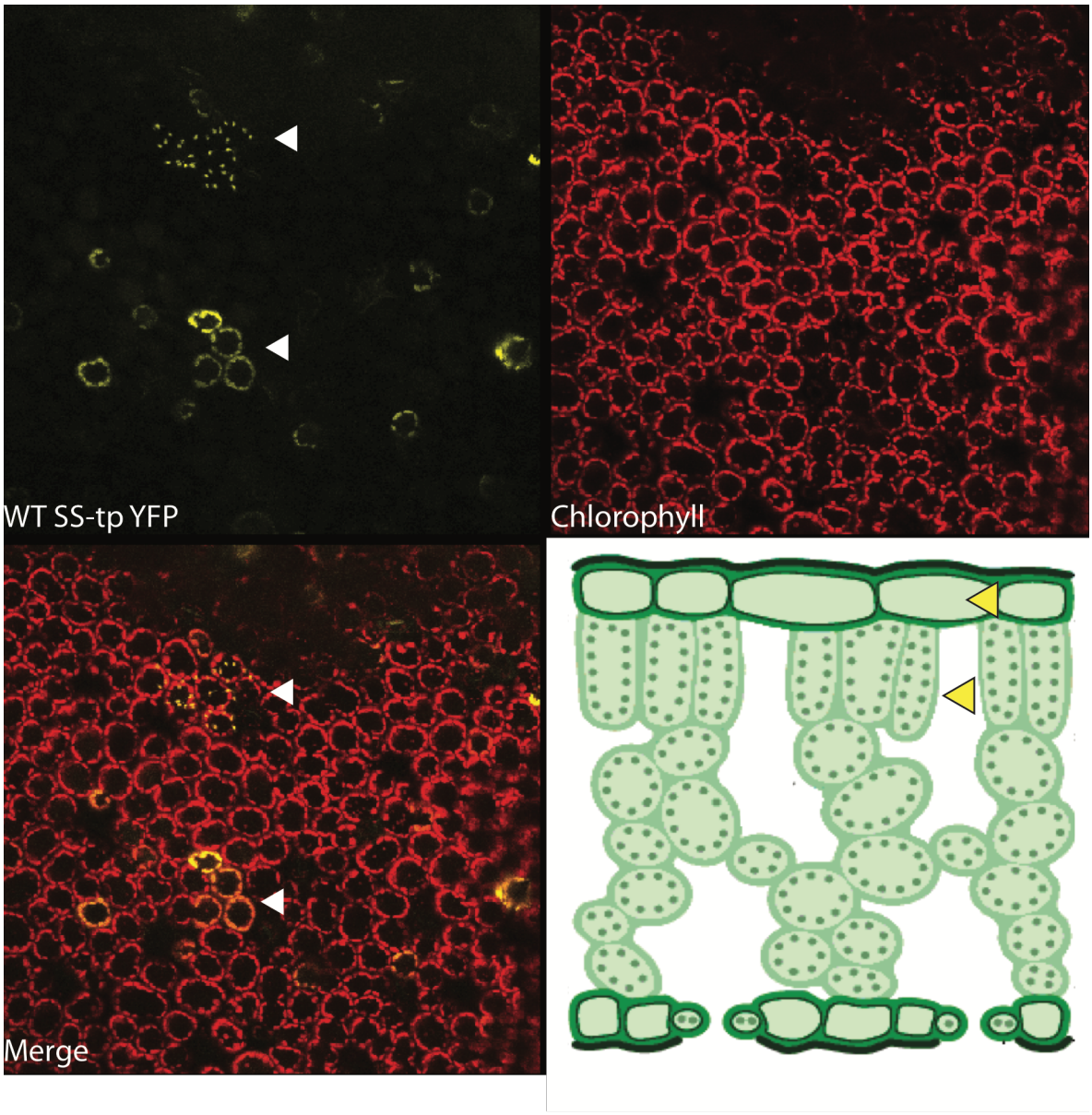


Figure 6-4 Continued

Although some variation in localization is expected, we find that the WT localization between sections is relatively consistent (Figure 6-7B). Next, we tested the timeframe of plastid protein localization. For this analysis, we selected M4 because this section represents a “middle” domain that is expected to have import attributes between the fully differentiated plastids in the “top” section and the undifferentiated leukoplasts in the “bottom” section. Starting at 2 hours post biolistic bombardment, we assessed as many cells as possible for fluorescence, in a section approximately 1 cm x 1 cm. Starting about about 6 hours post transformation, many transformed cells could be identified. Fluorescence localization was still easy to identify at 30 hrs post-bombardment (Figure 6-7C). Thus, leek biolistics represents a flexible, simple to quantify transformation method for plastid localization.

6.2.5 Leek plastids have a more uniform structure and lack stromules

Previous research has identified that both abiotic and biotic stressors can contribute to the formation of stromules. Stromules are thin, stroma-filled tubules that extend from the surface of all plastid types. These protrusions are highly dynamic, but the primary function of stromules remains unresolved (Natesan et al., 2005). Because stromules extend into the cytosol surrounding the plastid, their presence can cause artifacts in our quantification methodology. We investigated the ultrastructure of leek and onion plastids transiently expressing SStpNT-YFP using confocal microscopy. We find that onion plastids show a consistently higher occurrence of stromules as compared to leek plastids (Figure 6-8, compare top and bottom panels, stromules are denoted with a white arrow). Presence of stromules directly affects plastid versus cytosolic fluorescence ratios, as depicted in Figure 6-8B. Both the leek and onion samples were transformed with the same WT construct, but including stromules in the quantification introduces an artifactually lower plastid:cytosol fluorescence ratio in the onion sample. We are currently working on protocols to further assess leek and onion ultrastructure using transmission electron microscopy.

6.3 Discussion

Plastid protein import has been a focus of research for almost 40 years. By combining the *in organello* assays with *in vivo* microscopy, we can get a clearer picture of the factors

that influence precursor recognition and import. *In organello* assays use isolated pea chloroplasts and purified precursor protein. However, the ability to work with pea chloroplast in *in vivo* research is limited by the presence of chlorophyll. In previous work, we have utilized onion epidermal leukoplasts, which are amenable to our quantification method described in Chapter 4 (Chotewutmontri and Bruce, 2015; Chotewutmontri et al., 2012; Holbrook et al., 2016). Furthermore, some protein import studies utilize protoplasts derived from *Arabidopsis thaliana* (Lee et al., 2006; Lee, 2009). However, by developing protocols to use pea chloroplasts with *in vivo* microscopy, we would be able to better combine our *in vivo* and *in organello* work. We can also use these protocols in the future to compare the efficacy of different quantification methods. In this methods study, we first developed a pea protocol to efficiently produce protoplasts based on methods derived from Zhang et al. (Zhang, 2011). We find that several factors influence the production and quality of protoplasts, including the amount of plant tissue used initially and type of enzymes used. Removal of the epidermal layer appears to be critical. We also worked on improving methodology for the transient expression of SStp-YFP in pea seedlings using *Agrobacterium tumefaciens*. We find that using two rounds of vacuum infiltration greatly increases transformation rate. We observed transformation in both epidermal cells (punctate labeling) as well as mesophyll cells (round cells) using *Agrobacterium* (compare top and bottom panels of Figure 6-4).

Figure 6-5: Biolistic transformation of commercial leeks is more efficient than biolistic transformation of onions

A. Image depicts the structure of a commercially purchased leek. Dashed lines denote the approximate sectioning points for Bottom (B), Middle (M) and Top (T) sections for biolistic transformation, discussed later in this work. B-C. Transient expression of chimeric SStpNt-YFP constructs in leek and onion were live-imaged using epi-fluorescence microscopy. Top panel depicts GFP fluorescence and phase contrast representative images of transformed leek cells; C panels depict transformed onion cells. Scale bar denotes 1400 microns. C. Quantification of the number of transformed cells (positive GFP staining) for 5 images (150,000 microns^2 each). *** indicates conditions with a significantly higher transformation efficiency ($p<0.0001$) Unpaired t-test

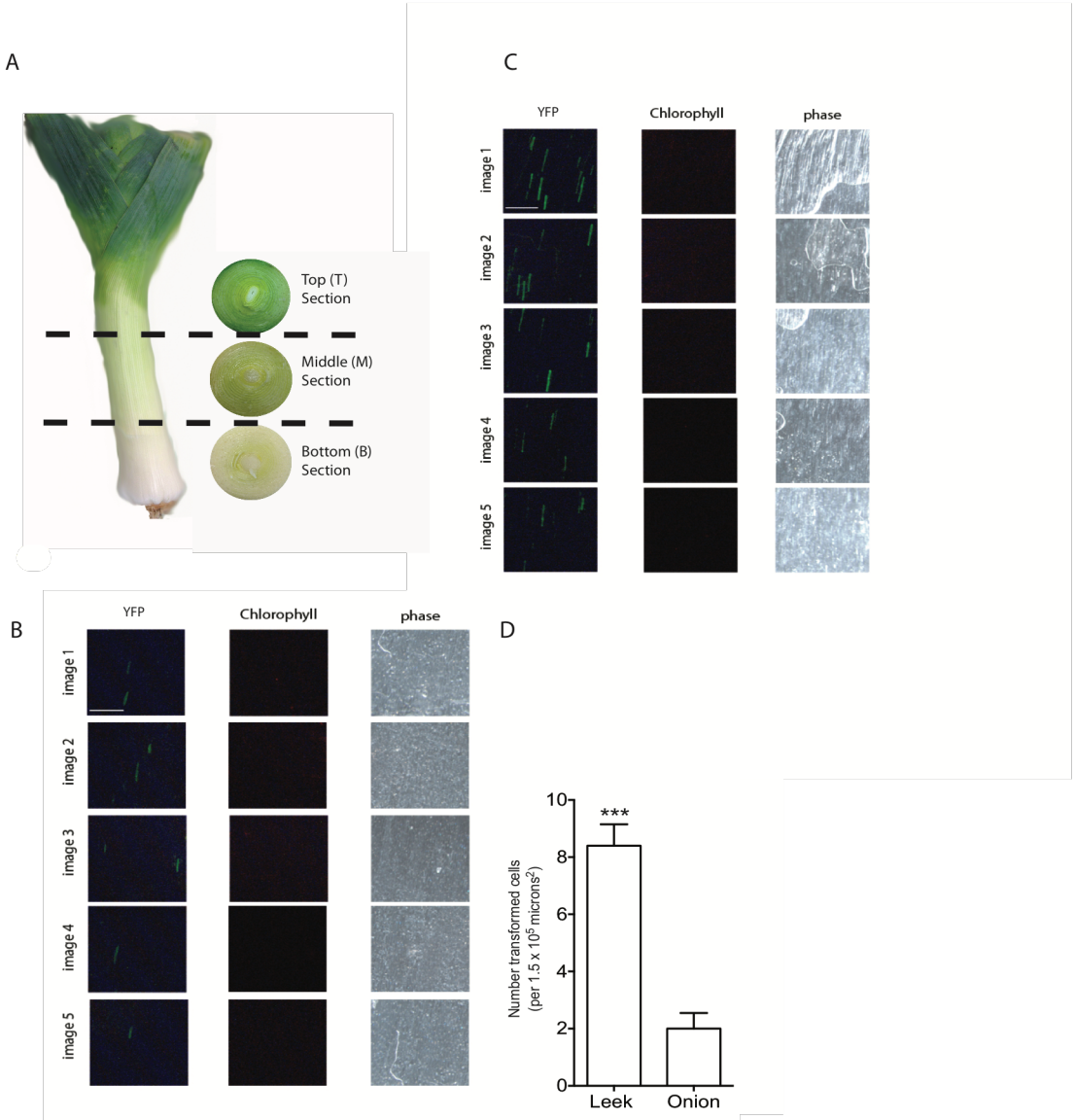


Figure 6-5 Continued

Figure 6-6: Chlorophyll and protein content in leek sections

A. Chlorophyll content approximation based on fluorescence in leek sections. Sections 4-5 correspond to “Top” region of the leek; section 3 corresponds to the “Middle” section, and sections 1-2 correspond to the “bottom” section. B. RIPA solubilized sections (corresponding to the sections in A) were analyzed using BCA analysis (Thermo fisher) for protein content. Measured protein content/mg of fresh weight (y-axis) was graphed by leek section number (x- axis)

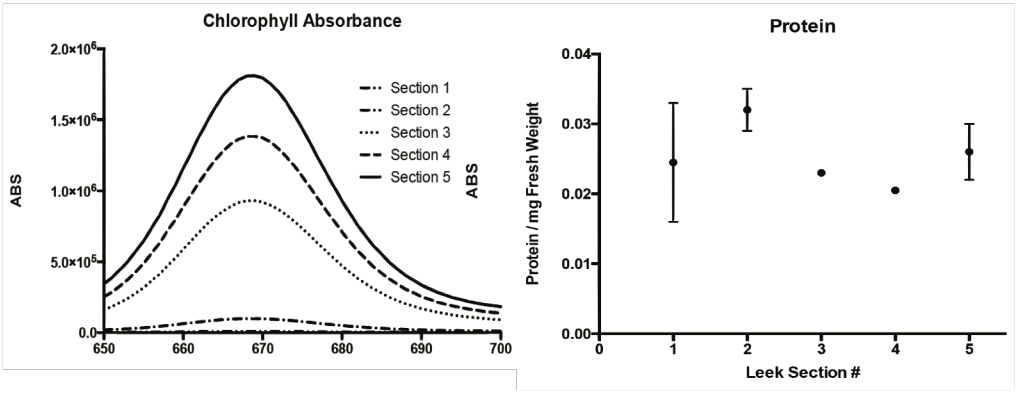


Figure 6-6 Continued

Figure 6-7: *In vivo* localization analysis of WT SStp-YFP in leeks

A. Image depicts approximate location of leek sections used for this study. 3,4,5 denote the section location within the leek bulb. B,M,T denote "bottom," "middle," or "top," respectively. See figure 6-2A. B. Points show the ratio of plastid localized signal intensity/cytosolic localized signal intensity 8 hours post biolistic transformation of commercial leek sections. B, M, and T represent the bottom, middle and top sections of the leek as defined in A. C. Points show the ratio of plastid localized signal intensity/cytosolic localized signal intensity post biolistic transformation. The number of hours post transformation is denoted on the X-axis. The results reflect the approximate signal intensity (YFP protein signal) of the SStpNT-YFP construct over time.

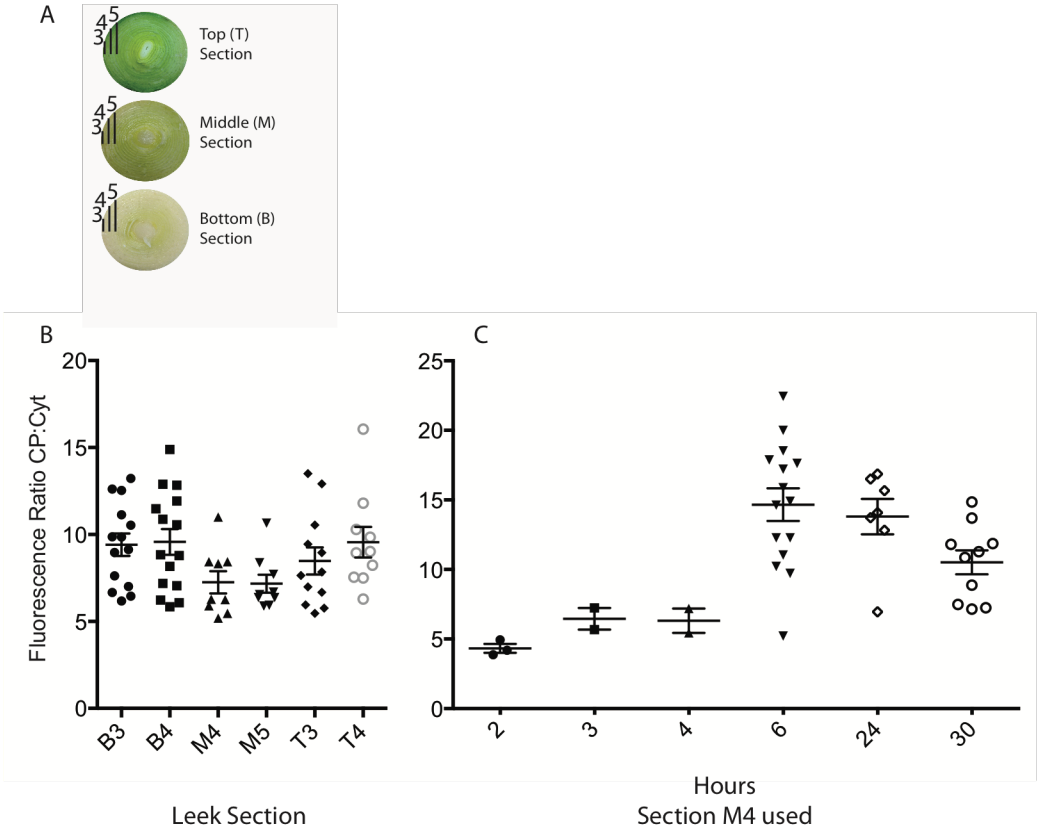
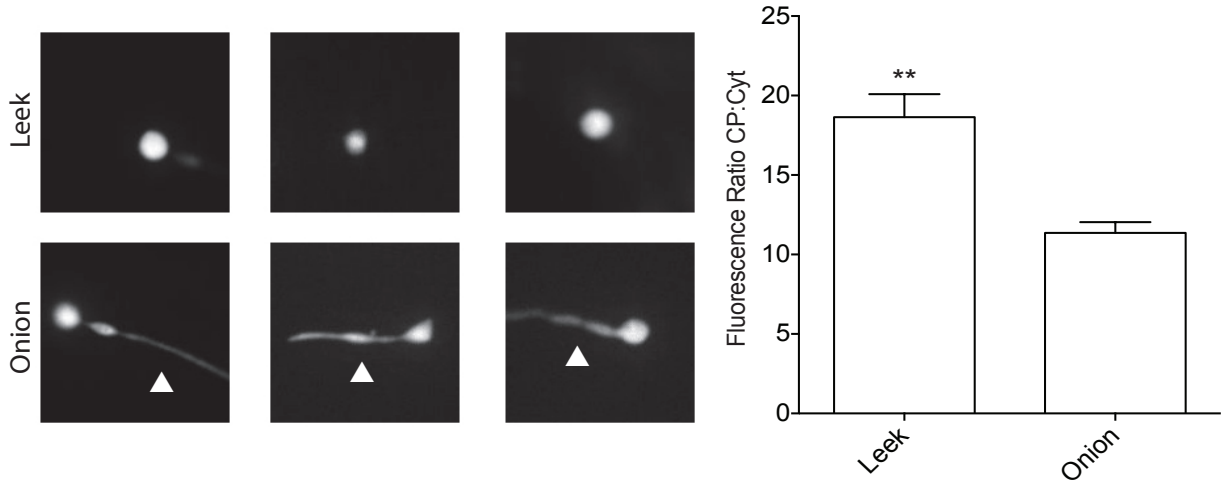


Figure 6-7 Continued

Figure 6-8: Confocal microscopy reveals the presence of stromules on onion plastids, but not leek plastids

A. Transient expression of chimeric SStpNt-YFP constructs in leek and onion were live-imaged using confocal microscopy. Top panel depicts GFP fluorescence of 3 representative images of transformed leek cells; bottom panels depict transformed onion cells. White arrows indicate stromules present on onion plastids. B. Bars show the ratio of plastid localized signal intensity/cytosolic localized signal intensity. ** indicate cells where WT SStpNt-YFP displays a significantly higher Plastid/Cytosol ratio (more plastid labeling) than cells where SStpNt-YFP is more localized in the cytosol. (N= 6 cells per condition, $p<0.001$) Unpaired T-test


Leek
Onion
Fluorescence Ratio CP:Cyt
25
20
15
10
5
0
**
Leek
Onion

Next, we will develop methodology to quantify plastid versus cytosolic expression of TP-YFP constructs in chloroplasts. Our analyses are made possible by improving imaging technology. We have recent access to the Leica SP8 LightGate microscope, new laser scanning confocal microscopy (LSCM) technology that greatly improves our ability to do complicated imaging. Previously, we were technically limited to analysis of protein import in leukoplasts due to the interference of high chlorophyll autofluorescence in green differentiated chloroplasts versus the fluorescence of TP-YFP. LightGate allows us to filter out reflected light based on fluorescence lifetime, and thus detected signal is purely fluorescence emission. We can use this technology to greatly improve our understanding of protein import in differentiated chloroplasts; this methodology will also allow us to compare similarities and differences between import in different plastid types. Although differentiated chloroplasts have been used for many years in *in organello* analysis, it is clear that combining *in vivo* physiological data will allow us to see a clearer picture of plastid import (compare Figure 4-2 and 4-7).

We have also worked on improving biolistic transformation protocols in the lab. Unexpectedly, we find that use of commercial leeks instead of onions resulted in better transformation efficiency. Leeks are also useful because a gradient of plastid differentiation is available (Figure 6-5A). Furthermore, leeks appear to have a consistently lower rate of stromule formation. In previous studies, stromules have been hypothesized to be associated with nutrient recycling under stress conditions (Mathur et al., 2013; Natesan et al., 2005). Because we are using commercially derived produce, it is difficult to predict the presence of stromules. However, a lower rate of stromules makes accurate protein localization quantification much easier. Future biolistic research in the Bruce lab will focus on improving existing onion based biolistic protocols for use with leeks. Ultimately, our goal is combine our knowledge of protein import assays with new microscopy technology to move our trafficking research forward.

CHAPTER 7: CONCLUSIONS AND FUTURE DIRECTIONS

A great deal of research focus has been dedicated to plastid protein import. The general import pathway balances the import of thousands of proteins while maintaining a high degree of specificity; this system is critical for organelle function. Our current research study has investigated several aspects of protein import; we focused on understanding both the physicochemical recognition elements of the TP as well as the mechanistic function of the TOC34 receptor component. The semi conserved FGLK domain appears to represent a central recognition element for the TP. Future work in the Bruce Lab will focus on the contributions of plastid OM membrane lipids and how they affect TP physicochemical elements.

Why do *in organello* and *in vivo* import results show some differences (compare Figure 7-1 and Figure 7-2)? The heat maps shown in Figure 7-1 and 7-2 summarize our findings using *in organello* versus *in vivo* analyses; the X-axis correlates to the TP region tested and the Y axis shows the mutant proteins we tested. Magenta color denotes a strong abrogation of WT TP recognition or import. The difference between *in organello* and *in vivo* assays is a complex question that appears to involve the presence of cytosolic factors. Although cytosolic factors were shown to associate with some precursors and in certain cases increase import efficiencies, these factors were found to be non-essential both *in organello* (Dabney-Smith et al., 1999; Pilon et al., 1990) and *in vivo* (Aronsson et al., 2007; Hofmann and Theg, 2005c; Nakrieko et al., 2004). The rate of *in organello* protein import without cytosolic factors was also suggested to be sufficient to sustain chloroplast development (Pilon et al., 1992d). Our study (2016) combined detailed analysis of TP import using *in organello* and *in vivo* assays. *In organello* analyses provide a highly quantitative and resolved biochemical insight while *in vivo* assays confirm a clear physiological role in living cells. However, it is clear that there are differential results between the *in organello* and *in vivo* results. Although mutagenesis of a single FGLK physicochemical domain shows a clear import defect in *in vitro* assays, mutagenesis of both domains is required to abrogate *in vivo* import. We hypothesize that cytosolic factors, which may somewhat aid in WT precursor targeting, are essential for the targeting of mutant proteins. To test this hypothesis, we plan to add cytosolic extract to our *in organello* assays to attempt to recapitulate our *in vivo* results.

The TOC34/TOC159 receptor components represent a translocon “lock” and their unique mechanistic role in import requires further study. We have developed a strong basis for understanding plastid import, future detailed analysis of the components mediating cytosolic recognition, TOC assembly, and regulation of the import process will farther develop our understanding of plastid import. Furthermore, understanding how diverse TPs mediate the selective targeting and translocation of precursors will reveal important regulation and quality control elements of this pathway. The integration and balance of many different import pathways for a single organelle exemplifies the complexity of eukaryotic cells. In the future we plan to integrate our *in vitro*, *in organello*, and emerging methodologies for *in vivo* assays to develop a clearer picture of the plastid import cycle. For example, the strong importance of TP basic residues (Chapter 4) may be related to the TOC34 GTP hydrolysis requirement for R133 (Chapter 3) (Figure 7-3). In summary, emerging studies that integrate the assembly and regulation of the plastid import pathways will contribute to uncovering the mechanisms that mediate the plasticity of organelle development.

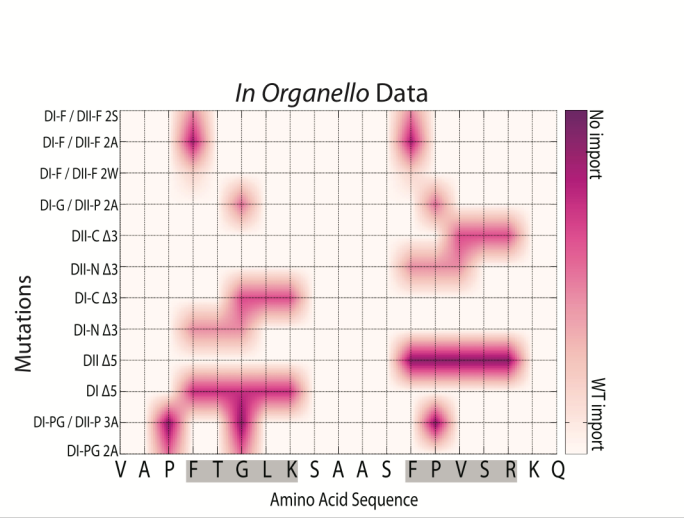


Figure 7-1: *In organello* heat map

Heat map summarizes *in organello* experiments described in Chapter 4. The X axis corresponds to the TP sequence; Y axis shows mutants tested in these experiments. Magenta color corrlates to a strong abrogation of WT TP recognition or import.

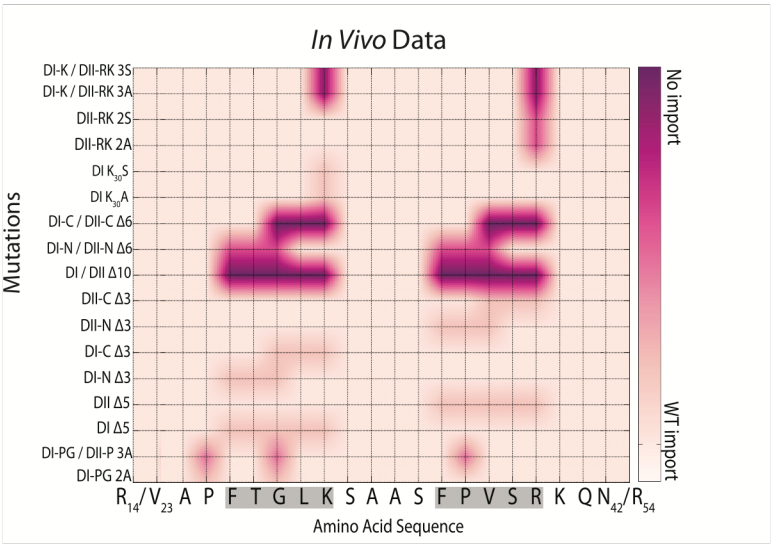


Figure 7-2: *In vivo* heat map

Heat map summarizes *in organello* experiments described in Chapter 4. The X axis corresponds to the TP sequence; Y axis shows mutants tested in these experiments. Magenta color corrlates to a strong abrogation of WT TP recognition or import.

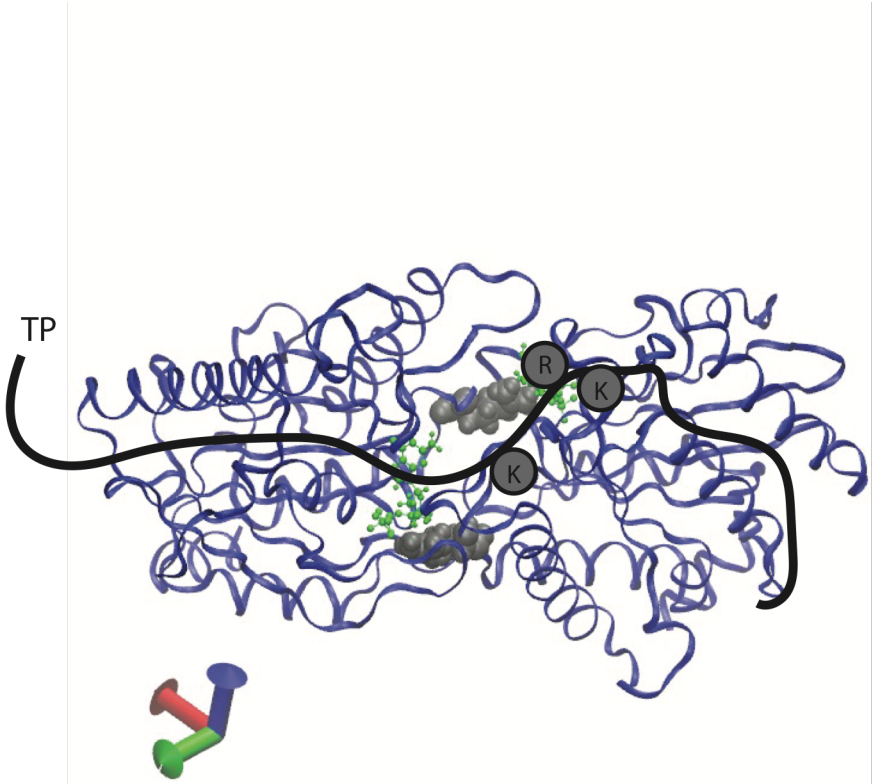


Figure 7-3: Potential requirement for positive charge in TP recognition
Cartoon schematic showing hypothesized interaction between TP and TOC34, highlighting the importance of both TP basic residues and the importance of R133 in the TOC34 GTPase cycle. GDP is shown in green ball-and-stick format. TOC34 crystal structure is shown as a blue ribbon diagram; R133 is shown as a gray-spacing filling model.

LIST OF REFERENCES

Abramoff, M.D., Magelhaes, P.J., and Ram, S.J. (2004). Image Processing with ImageJ. Biophotonics Int *11*, 36-42.
Akita, M., and Inoue, H. (2009). Evaluating the energy-dependent "binding" in the early stage of protein import into chloroplasts. Methods Enzymol *466*, 43-64.
Akita, M., Nielsen, E., and Keegstra, K. (1997). Identification of protein transport complexes in the chloroplastic envelope membranes via chemical cross-linking. J Cell Biol *136*, 983-994.
Aleshin, A.E., Gramatikova, S., Hura, G.L., Bobkov, A., Strongin, A.Y., Stec, B., Tainer, J.A., Liddington, R.C., and Smith, J.W. (2009). Crystal and solution structures of a prokaryotic M16B peptidase: an open and shut case. Structure *17*, 1465-1475.
Allen, J.F. (2003). The function of genomes in bioenergetic organelles. Philosophical transactions of the Royal Society of London Series B, Biological sciences *358*, 19-37; discussion 37-18.
America, T., Hageman, J., Guera, A., Rook, F., Archer, K., Keegstra, K., and Weisbeek, P. (1994). Methotrexate does not block import of a DHFR fusion protein into chloroplasts. Plant Mol Biol *24*, 283-294.
Andersson, M.X., and Sandelius, A.S. (2004). A chloroplast-localized vesicular transport system: a bio-informatics approach. BMC Genomics *5*, 40-40.
Andreotti, A. (2003). Native state proline isomerization: an intrinsic molecular switch. Biochemistry *42*, 9515-9524.
Andrès, C., Agne, B., and Kessler, F. (2010). The TOC complex: Preprotein gateway to the chloroplast. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1803*, 715-723.
Aronsson, H. (2008). The galactolipid monogalactosyldiacylglycerol (MGDG) contributes to photosynthesis-related processes in Arabidopsis thaliana. Plant Signal Behav *3*, 1093-1095.
Aronsson, H., Combe, J., Patel, R., Agne, B., Martin, M., Kessler, F., and 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 *63*, 297-311.
Balsera, M., Goetze, T.A., Kovacs-Bogdan, E., Schurmann, P., Wagner, R., Buchanan, B.B., Soll, J., and Bolter, B. (2009a). Characterization of Tic110, a channel-forming protein at the inner envelope membrane of chloroplasts, unveils a response to Ca(2+) and a stromal regulatory disulfide bridge. J Biol Chem *284*, 2603-2616.
Balsera, M., Soll, J., and Bolter, B. (2009b). Protein import machineries in endosymbiotic organelles. Cell Mol Life Sci *66*, 1903-1923.
Balsera, M., Soll, J., and Buchanan, B.B. (2010). Redox extends its regulatory reach to chloroplast protein import. Trends Plant Sci *15*, 515-521.
Bang, W.Y., Chen, J., Jeong, I.S., Kim, S.W., Kim, C.W., Jung, H.S., Lee, K.H., Kweon, H.-S., Yoko, I., Shiina, T.*, et al.* (2012). Functional characterization of ObgC in ribosome biogenesis during chloroplast development. The Plant Journal *71*, 122-134.
Baudisch, B., Langner, U., Garz, I., and Klösgen, R.B. (2014). The exception proves the rule? Dual targeting of nuclear-encoded proteins into endosymbiotic organelles. New Phytologist *201*, 80-90.

Bauer, J., Hiltbrunner, A., Weibel, P., Vidi, P.A., Alvarez-Huerta, M., Smith, M.D., Schnell, D.J., and Kessler, F. (2002). Essential role of the G-domain in targeting of the protein import receptor atToc159 to the chloroplast outer membrane. J Cell Biol *159*, 845-854.
Baum, D. (2013). The Origin of Primary Plastids: A Pas de Deux or a Menage a Trois? Plant Cell.
Becker, T., Jelic, M., Vojta, A., Radunz, A., Soll, J., and Schleiff, E. (2004). Preprotein recognition by the Toc complex. The EMBO journal *23*, 520-530.
Bedard, J., and Jarvis, P. (2005). Recognition and envelope translocation of chloroplast preproteins. J Exp Bot *56*, 2287-2320.
Bhushan, S., Kuhn, C., Berglund, A.-K., Roth, C., and Glaser, E. (2006). The role of the N-terminal domain of chloroplast targeting peptides in organellar protein import and miss-sorting. FEBS Letters *580*, 3966-3972.
Bhushan, S., Lefebvre, B., Stahl, A., Wright, S.J., Bruce, B.D., Boutry, M., and Glaser, E. (2003). Dual targeting and function of a protease in mitochondria and chloroplasts. EMBO reports *4*, 1073-1078.
Bhushan, S., Stahl, A., Nilsson, S., Lefebvre, B., Seki, M., Roth, C., McWilliam, D., Wright, S.J., Liberles, D.A., Shinozaki, K.*, et al.* (2005). Catalysis, subcellular localization, expression and evolution of the targeting peptides degrading protease, AtPreP2. Plant & cell physiology *46*, 985-996.
Bian, X., Klemm, R.W., Liu, T.Y., Zhang, M., Sun, S., Sui, X., Liu, X., Rapoport, T.A., and Hu, J. (2011). Structures of the atlastin GTPase provide insight into homotypic fusion of endoplasmic reticulum membranes. Proceedings of the National Academy of Sciences of the United States of America *108*, 3976-3981.
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. Journal of Molecular Biology *402*, 510-523.
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. Journal of Molecular Biology *402*, 510-523.
Blobel, G. (1980). Intracellular protein topogenesis. Proc Natl Acad Sci U S A *77*, 1496-1500.
Bos, J.L., Rehmann, H., and Wittinghofer, A. (2007). GEFs and GAPs: critical elements in the control of small G proteins. Cell *129*, 865-877.
Bourne, H.R., Sanders, D.A., and McCormick, F. (1991). The GTPase superfamily: conserved structure and molecular mechanism. Nature *349*, 117-127.
Brandizzi, F. (2011). Is There a COPII-Mediated Membrane Traffic in Chloroplasts? Traffic *12*, 9-11.
Bruce, B., Perry, S., Froehlich, J., and Keegstra, K. (1994). In vitro import of proteins into chloroplasts. In Plant Molecular Biology Manual, S. Gelvin, and R. Schilperoort, eds. (Springer Netherlands), pp. 497-511.
Bruce, B.D. (1998a). The role of lipids in plastid protein transport. Plant Mol Biol *38*, 223-246.
Bruce, B.D. (1998b). The role of lipids in plastid protein transport. . Plant Mol Biol *38*, 223-246.

Bruce, B.D. (2000a). Chloroplast transit peptides: structure, function and evolution. Trends in cell biology *10*, 440-447.
Bruce, B.D. (2000b). Chloroplast transit peptides: structure, function and evolution. . Trends Cell Biol *10*.
Bruce, B.D. (2001a). The paradox of plastid transit peptides: conservation of function despite divergence in primary structure. Biochim Biophys Acta *1541*, 2-21.
Bruce, B.D. (2001b). The paradox of plastid transit peptides: conservation of function despite divergence in primary structure. Biochim Biophys Acta *1541*, 2-21.
Butterfield, N.J. (2000). Bangiomorpha pubescens n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. Paleobiology *26*, 386-404.
Caliebe, A., Grimm, R., Kaiser, G., Lubeck, J., Soll, J., and Heins, L. (1997). The chloroplastic protein import machinery contains a Rieske-type iron-sulfur cluster and a mononuclear iron-binding protein. EMBO J *16*, 7342-7350.
Carrie, C., Giraud, E., and Whelan, J. (2009). Protein transport in organelles: Dual targeting of proteins to mitochondria and chloroplasts. FEBS Journal *276*, 1187-1195.
Carrie, C., and Small, I. (2013). A reevaluation of dual-targeting of proteins to mitochondria and chloroplasts. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1833*, 253-259.
Carrie, C., and Whelan, J. (2013). Widespread dual targeting of proteins in land plants: When, where, how and why. Plant Signaling & Behavior *8*, e25034.
Chappie, J.S., Acharya, S., Leonard, M., Schmid, S.L., and Dyda, F. (2010). G domain dimerization controls dynamin's assembly-stimulated GTPase activity. Nature *465*, 435-440.
Chen, K.-Y., and Li, H.-m. (2007a). Precursor binding to an 880-kDa Toc complex as an early step during active import of protein into chloroplasts. The Plant Journal *49*, 149-158.
Chen, K.Y., and Li, H.M. (2007b). Precursor binding to an 880-kDa Toc complex as an early step during active import of protein into chloroplasts. The Plant journal : for cell and molecular biology *49*, 149-158.
Chen, L.J., and Li, H.M. (1998). A mutant deficient in the plastid lipid DGD is defective in protein import into chloroplasts. Plant J *16*, 33-39.
Chen, X., Smith, M.D., Fitzpatrick, L., and Schnell, D.J. (2002). In vivo analysis of the role of atTic20 in protein import into chloroplasts. Plant Cell *14*, 641-654.
Chen, Y.-B., Lu, T.-C., Wang, H.-X., Shen, J., Bu, T.-T., Chao, Q., Gao, Z.-F., Zhu, X.-G., Wang, Y.-F., and Wang, B.-C. (2014). Posttranslational modification of maize chloroplast pyruvate orthophosphate dikinase reveals the precise regulatory mechanism of its enzymatic activity. Plant Physiology *165*, 534-549.
Chigri, F., Hormann, F., Stamp, A., Stammers, D.K., Bolter, B., Soll, J., and Vothknecht, U.C. (2006). Calcium regulation of chloroplast protein translocation is mediated by calmodulin binding to Tic32. Proc Natl Acad Sci U S A *103*, 16051-16056.
Chirico, W.J., Waters, M.G., and Blobel, G. (1988). 70K heat shock related proteins stimulate protein translocation into microsomes. Nature *332*, 805-810.
Chotewutmontri, P., and Bruce, B.D. (2015). Non-native, N-terminal Hsp70 molecular motor recognition elements in transit peptides support plastid protein translocation. Journal of Biological Chemistry *290*, 7602-7621.

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. The Plant Cell *24*, 3040-3059.
Chou, M.L., Chu, C.C., Chen, L.J., Akita, M., and Li, H.M. (2006). Stimulation of transit-peptide release and ATP hydrolysis by a cochaperone during protein import into chloroplasts. The Journal of cell biology *175*, 893-900.
Chou, M.L., Fitzpatrick, L.M., Tu, S.L., Budziszewski, G., Potter-Lewis, S., Akita, M., Levin, J.Z., Keegstra, K., and Li, H.M. (2003). Tic40, a membrane-anchored co-chaperone homolog in the chloroplast protein translocon. EMBO J *22*, 2970-2980.
Chu, C.-C., and Li, H.-m. (2015). Protein import into isolated pea root leucoplasts. Frontiers in Plant Science *6*, 690.
Chua, N.H., and Schmidt, G.W. (1979). Transport of Proteins into Mitochondria and Chloroplasts. Journal of Cell Biology *81*, 461-483.
Chupin, V., van 't Hof, R., and de Kruijff, B. (1994). The transit sequence of a chloroplast precursor protein reorients the lipids in monogalactosyl diglyceride containing bilayers. FEBS Lett *350*, 104-108.
Cleary, S.P., Tan, F.-C., Nakrieko, K.-A., Thompson, S.J., Mullineaux, P.M., Creissen, G.P., von Stedingk, E., Glaser, E., Smith, A.G., and Robinson, C. (2002). Isolated Plant Mitochondria Import Chloroplast Precursor Proteinsin Vitro with the Same Efficiency as Chloroplasts. Journal of Biological Chemistry *277*, 5562-5569.
Cline, K., and Henry, R. (1996). Import and routing of nucleus-encoded chloroplast proteins. Annual review of cell and developmental biology *12*, 1-26.
Cline, K., Keegstra, K., and Staehelin, L.A. (1985). Freeze-fracture electron microscopic analysis of ultrarapidly frozen envelope membranes on intact chloroplasts and after purification. Protoplasma *125*, 111-123.
Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method forAgrobacterium-mediated transformation ofArabidopsis thaliana. The Plant Journal *16*, 735-743.
Constan, D., Froehlich, J.E., Rangarajan, S., and Keegstra, K. (2004). A stromal Hsp100 protein is required for normal chloroplast development and function in Arabidopsis. Plant Physiology *136*, 3605-3615.
Criscuolo, A., and Gribaldo, S. (2011). Large-scale phylogenomic analyses indicate a deep origin of primary plastids within cyanobacteria. Molecular biology and evolution *28*, 3019-3032.
Dabney-Smith, C., van Den Wijngaard, P.W., Treece, Y., Vredenberg, W.J., and Bruce, B.D. (1999). The C terminus of a chloroplast precursor modulates its interaction with the translocation apparatus and PIRAC. The Journal of biological chemistry *274*, 32351-32359.
Dabney-Smith, C., van Den Wijngaard, P.W., Treece, Y., Vredenberg, W.J., and Bruce, B.D. (1999). The C terminus of a chloroplast precursor modulates its interaction with the translocation apparatus and PIRAC. J Biol Chem *274*, 32351-32359.
Dahlin, C., and Cline, K. (1991). Developmental regulation of the plastid protein import apparatus. Plant Cell *3*, 1131-1140.
Danpure, C.J. (1995). How can the products of a single gene be localized to more than one intracellular compartment? Trends in Cell Biology *5*, 230-238.

de Castro Silva Filho, M., Chaumont, F., Leterme, S., and Boutry, M. (1996). Mitochondrial and chloroplast targeting sequences in tandem modify protein import specificity in plant organelles. . Plant Mol Biol *30*, 769-780.
Deshaies, R.J., Koch, B.D., Werner-Washburne, M., Craig, E.A., and Schekman, R. (1988). A subfamily of stress proteins facilitates translocation of secretory and mitochondrial precursor polypeptides. Nature *332*, 800-805.
Dhanoa, P.K., Richardson, L.G.L., Smith, M.D., Gidda, S.K., Henderson, M.P.A., Andrews, D.W., and Mullen, R.T. (2010). Distinct Pathways Mediate the Sorting of Tail-Anchored Proteins to the Plastid Outer Envelope. PLoS ONE *5*, e10098.
Dobberstein, B., Blobel, G., and Chua, N.H. (1977). In vitro synthesis and processing of a putative precursor for the small subunit of ribulose-1,5-bisphosphate carboxylase of Chlamydomonas reinhardtii. Proceedings of the National Academy of Sciences of the United States of America *74*, 1082-1085.
Dobberstein, B., Blobel, G., and Chua, N.H. (1977). In vitro synthesis and processing of a putative precursor for the small subunit of ribulose-1,5-bisphosphate carboxylase of Chlamydomonas reinhardtii. . Proc Natl Acad Sci U S A *74*, 1082-1085.
Douce, R., and Joyard, J. (1990). Biochemistry and Function of the Plastid Envelope. Annual review of cell biology *6*, 173-216.
Douglas, S.E. (1998). Plastid evolution: origins, diversity, trends. Current opinion in genetics & development *8*, 655-661.
Drea, S.C., Mould, R.M., Hibberd, J.M., Gray, J.C., and Kavanagh, T.A. (2001). Tissue-specific and developmental-specific expression of an Arabidopsis thaliana gene encoding the lipoamide dehydrogenase component of the plastid pyruvate dehydrogenase complex. Plant Mol Biol *46*, 705-715.
Dugave, C., and Demange, L. (2003). Cis−Trans Isomerization of Organic Molecules and Biomolecules: Implications and Applications†. Chemical Reviews *103*, 2475-2532.
Emanuelsson, O., Nielsen, H., and von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci *8*, 978-984.
Estavillo, G.M., Crisp, P.A., Pornsiriwong, W., Wirtz, M., Collinge, D., Carrie, C., Giraud, E., Whelan, J., David, P., Javot, H.*, et al.* (2011). Evidence for a SAL1-PAP Chloroplast Retrograde Pathway That Functions in Drought and High Light Signaling in Arabidopsis. Plant Cell *23*, 3992-4012.
Fellerer, C., Schweiger, R., Schöngruber, K., Soll, J., and Schwenkert, S. (2011). Cytosolic HSP90 Cochaperones HOP and FKBP Interact with Freshly Synthesized Chloroplast Preproteins of Arabidopsis. Molecular plant *4*, 1133-1145.
Flores-Pérez, Ú., Bédard, J., Tanabe, N., Lymperopoulos, P., Clarke, A.K., and Jarvis, P. (2016). Functional Analysis of the Hsp93/ClpC Chaperone at the Chloroplast Envelope. Plant Physiology *170*, 147-162.
Friedman, A.L., and Keegstra, K. (1989). Chloroplast Protein Import : Quantitative Analysis of Precursor Binding. Plant Physiol *89*, 993-999.
Gagat, P., Bodył, 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. Biology Direct *8*, 1-22.

Gasper, R., Meyer, S., Gotthardt, K., Sirajuddin, M., and Wittinghofer, A. (2009). It takes two to tango: regulation of G proteins by dimerization. . Nat Rev Mol Cell Biol *10*, 423-429.
Gavel, Y., and von Heijne, G. (1990). A conserved cleavage-site motif in chloroplast transit peptides. FEBS Lett *261*, 455-458.
Ge, C., Spanning, E., Glaser, E., and Wiesland, A. (2014). Import determinants of organelle-specific and dual targeting peptides of mitochondria and chloroplasts in Arabidopsis thaliana. Molecular plant *7*, 121-136.
Gesch, R.W., Kang, I.H., Gallo-Meagher, M., Vu, J.C.V., Boote, K.J., Allen, L.H., and Bowes, G. (2003). Rubisco expression in rice leaves is related to genotypic variation of photosynthesis under elevated growth CO2 and temperature. Plant Cell Environ *26*, 1941-1950.
Glaser, E., Nilsson, S., and Bhushan, S. (2006). Two novel mitochondrial and chloroplastic targeting-peptide-degrading peptidasomes in A. thaliana, AtPreP1 and AtPreP2. Biol Chem *387*, 1441-1447.
Gouveia-Oliveira, R., Sackett, P., and Pedersen, A. (2007). MaxAlign: maximizing usable data in an alignment. BMC Bioinformatics *8*.
Gray, J.C., Sullivan, J.A., Wang, J.H., Jerome, C.A., and MacLean, D. (2003). Coordination of plastid and nuclear gene expression. Philos T R Soc B *358*, 135-144.
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. The Plant journal : for cell and molecular biology *23*, 771-783.
Guy, C.L., and Li, Q.B. (1998). The organization and evolution of the spinach stress 70 molecular chaperone gene family. Plant Cell *10*, 539-556.
Harmer, S.L., Hogenesch, J.B., Straume, M., Chang, H.S., Han, B., Zhu, T., Wang, X., Kreps, J.A., and Kay, S.A. (2000). Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. Science *290*, 2110-2113.
Heins, L., Mehrle, A., Hemmler, R., Wagner, R., Kuchler, M., Hormann, F., Sveshnikov, D., and Soll, J. (2002). The preprotein conducting channel at the inner envelope membrane of plastids. EMBO J *21*, 2616-2625.
Hiltbrunner, A., Bauer, J., Vidi, P.A., Infanger, S., Weibel, P., Hohwy, M., and Kessler, F. (2001). Targeting of an abundant cytosolic form of the protein import receptor at Toc159 to the outer chloroplast membrane. J Cell Biol *154*, 309-316.
Hinnah, S.C., Hill, K., Wagner, R., Schlicher, T., and Soll, J. (1997). Reconstitution of a chloroplast protein import channel. EMBO J *16*, 7351-7360.
Hirohashi, T., Hase, T., and Nakai, M. (2001). Maize Non-Photosynthetic Ferredoxin Precursor Is Mis-Sorted to the Intermembrane Space of Chloroplasts in the Presence of Light. Plant Physiology *125*, 2154-2163.
Hofmann, N.R., and Theg, S.M. (2005a). Chloroplast outer membrane protein targeting and insertion. Trends Plant Sci *10*, 450-457.
Hofmann, N.R., and Theg, S.M. (2005b). Protein- and energy-mediated targeting of chloroplast outer envelope membrane proteins. Plant J *44*, 917-927.
Holbrook, K., Subramanian, C., Chotewutmontri, P., Reddick, L.E., Wright, S.J., Zhang, H., Moncrief, L., and Bruce, B.D. (2016). Functional analysis of semi-conserved transit peptide motifs and mechanistic implications in precursor targeting and recognition. Molecular plant.

Huang, S., Taylor, N.L., Narsai, R., Eubel, H., Whelan, J., and Millar, A.H. (2009a). Experimental Analysis of the Rice Mitochondrial Proteome, Its Biogenesis, and Heterogeneity. Plant Physiology *149*, 719-734.
Huang, S., Taylor, N.L., Whelan, J., and Millar, A.H. (2009b). Refining the Definition of Plant Mitochondrial Presequences through Analysis of Sorting Signals, N-Terminal Modifications, and Cleavage Motifs. Plant Physiology *150*, 1272-1285.
Inaba, T., Alvarez-Huerta, M., Li, M., Bauer, J., Ewers, C., Kessler, F., and Schnell, D.J. (2005). Arabidopsis tic110 is essential for the assembly and function of the protein import machinery of plastids. Plant Cell *17*, 1482-1496.
Inoue, H., and Akita, M. (2008). Three sets of translocation intermediates are formed during the early stage of protein import into chloroplasts. J Biol Chem *283*, 7491-7502.
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.
Inoue, K. (2015). Emerging knowledge of the organelle outer membranes – research snapshots and an updated list of the chloroplast outer envelope proteins. Frontiers in Plant Science *6*, 278.
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.
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.
Ivey, R.A., 3rd, Subramanian, C., and Bruce, B.D. (2000a). Identification of a Hsp70 recognition domain within the rubisco small subunit transit peptide. Plant Physiol *122*, 1289-1299.
Ivey, R.A.I., Subramanian, C., and Bruce, B.D. (2000b). Identification of a HSP70 recognition domain within the rubisco small subunit transit peptide. Plant Physiology *122*, 1289-1299.
Jabs, A., Weiss, M.S., and Hilgenfeld, R. (1999). Non-proline cis peptide bonds in proteins. J Mol Biol *286*, 291-304.
Jackson-Constan, D., Akita, M., and Keegstra, K. (2001). Molecular chaperones involved in chloroplast protein import. Biochim Biophys Acta *1541*, 102-113.
Jarvis, P. (2008a). Targeting of nucleus-encoded proteins to chloroplasts in plants. New Phytol *179*, 257-285.
Jarvis, P. (2008b). Targeting of nucleus-encoded proteins to chloroplasts in plants. New Phytol *179*, 257-285.
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.
Jelic, M., Soll, J., and Schleiff, E. (2003). Two Toc34 homologues with different properties. Biochemistry *42*, 5906-5916.
Jelic, M., Sveshnikova, N., Motzkus, M., Hoerth, P., Soll, J., and Schleiff, E. (2002a). The chloroplast import receptor Toc34 functions as a preprotein regulated GTPase. Biological Chemistry *383*, 1875-1883.

Jelic, M., Sveshnikova, N., Motzkus, M., Horth, P., Soll, J., and Schleiff, E. (2002b). The chloroplast import receptor Toc34 functions as preprotein-regulated GTPase. Biological chemistry *383*, 1875-1883.
Joyard, J., Block, M.A., and Douce, R. (1991). Molecular aspects of plastid envelope biochemistry. Eur J Biochem *199*, 489-509.
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.
Karim, S., and Aronsson, H. (2014). The puzzle of chloroplast vesicle transport – involvement of GTPases. Frontiers in Plant Science *5*, 472.
Karlin-Neumann, G.A., and Tobin, E.M. (1986). Transit peptides of nuclear-encoded chloroplast proteins share a common amino acid framework. The EMBO journal *5*, 9-13.
Katoh, K., Kuma, K., Toh, H., and Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res *33*, 511-518.
Keegstra, K. (1989). Transport and routing of proteins into chloroplasts. Cell *56*, 247-253.
Keegstra, K., and Cline, K. (1999). Protein import and routing systems of chloroplasts. . Plant Cell *11*, 557-570.
Keegstra, K., and Cline, K. (1999). Protein import and routing systems of chloroplasts. The Plant cell *11*, 557-570.
Keller, R., and Schneider, D. (2013). Homologs of the yeast Tvp38 vesicle-associated protein are conserved in chloroplasts and cyanobacteria. Frontiers in Plant Science *4*, 467.
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.
Kessler, F., and Schnell, D.J. (2002). A GTPase gate for protein import into chloroplasts. Nat Struct Biol *9*, 81-83.
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.
Khan, N.Z., Lindquist, E., and Aronsson, H. (2013). New Putative Chloroplast Vesicle Transport Components and Cargo Proteins Revealed Using a Bioinformatics Approach: An Arabidopsis Model. PLoS ONE *8*, e59898.
Kikuchi, S., Bedard, J., Hirano, M., Hirabayashi, Y., Oishi, M., Imai, M., Takase, M., Ide, T., and Nakai, M. (2013). Uncovering the protein translocon at the chloroplast inner envelope membrane. Science *339*, 571-574.
Kikuchi, S., Hirohashi, T., and Nakai, M. (2006a). Characterization of the preprotein translocon at the outer envelope membrane of chloroplasts by blue native PAGE. Plant & cell physiology *47*, 363-371.
Kikuchi, S., Hirohashi, T., and Nakai, M. (2006b). Characterization of the Preprotein Translocon at the Outer Envelope Membrane of Chloroplasts by Blue Native PAGE. Plant and Cell Physiology *47*, 363-371.
Kilian, O., and Kroth, P.G. (2004). Presequence Acquisition During Secondary Endocytobiosis and thePossible Role of Introns. Journal of Molecular Evolution *58*, 712-721.
Kim, J.S., and Raines, R.T. (1993). Ribonuclease S-peptide as a carrier in fusion proteins. Protein Sci *2*, 348-356.

Kindle, K.L. (1998a). Amino-terminal and hydrophobic regions of the Chlamydomonas reinhardtii plastocyanin transit peptide are required for efficient protein accumulation in vivo. . Plant Mol Biol *38*, 365-377.
Kindle, K.L., and Lawrence, S.D. (1998b). Transit peptide mutations that impair in vitro and in vivo chloroplast protein import do not affect accumulation of the gamma-subunit of chloroplast ATPase. . Plant Physiology *116*, 1179-1190.
Kleffmann, T., Russenberger, D., von Zychlinski, A., Christopher, W., Sjolander, K., Gruissem, W., and Baginsky, S. (2004). The Arabidopsis thaliana chloroplast proteome reveals pathway abundance and novel protein functions. Curr Biol *14*, 354-362.
Knight, J.S., Duckett, C.M., Sullivan, J.A., Walker, A.R., and Gray, J.C. (2002). Tissue-specific, light-regulated and plastid-regulated expression of the single-copy nuclear gene encoding the chloroplast Rieske FeS protein of Arabidopsis thaliana. Plant & cell physiology *43*, 522-531.
Ko, K., Bornemisza, O., Kourtz, L., Ko, Z.W., Plaxton, W.C., and Cashmore, A.R. (1992). Isolation and Characterization of a Cdna Clone Encoding a Cognate 70-Kda Heat-Shock Protein of the Chloroplast Envelope. Journal of Biological Chemistry *267*, 2986-2993.
Koenig, P., Oreb, M., Hofle, A., Kaltofen, S., Rippe, K., Sinning, I., Schleiff, E., and Tews, I. (2008a). The GTPase cycle of the chloroplast import receptors Toc33/Toc34: implications from monomeric and dimeric structures. Structure *16*, 585-596.
Koenig, P., Oreb, M., Rippe, K., Muhle-Goll, C., Sinning, I., Schleiff, E., and Tews, I. (2008b). On the significance of Toc-GTPase homodimers. J Biol Chem *283*, 23104-23112.
Kouranov, A., Chen, X., Fuks, B., and Schnell, D.J. (1998). Tic20 and Tic22 are new components of the protein import apparatus at the chloroplast inner envelope membrane. J Cell Biol *143*, 991-1002.
Kouranov, A., and Schnell, D.J. (1997). Analysis of the interactions of preproteins with the import machinery over the course of protein import into chloroplasts. J Cell Biol *139*, 1677-1685.
Kourtz, L., and Ko, K. (1997). The early stage of chloroplast protein import involves Com70. Journal of Biological Chemistry *272*, 2808-2813.
Kovacheva, S., Bedard, J., Patel, R., Dudley, P., Twell, D., Rios, G., Koncz, C., and Jarvis, P. (2005). In vivo studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. Plant J *41*, 412-428.
Kovacheva, S., Bedard, J., Wardle, A., Patel, R., and Jarvis, P. (2007). Further in vivo studies on the role of the molecular chaperone, Hsp93, in plastid protein import. Plant Journal *50*, 364-379.
Krimm, I., Gans, P., Hernandez, J.F., Arlaud, G.J., and Lancelin, J.M. (1999). A coil-helix instead of a helix-coil motif can be induced in a chloroplast transit peptide from Chlamydomonas reinhardtii. Eur J Biochem *265*, 171-180.
Kubis, S., Patel, R., Combe, J., Bedard, J., Kovacheva, S., Lilley, K., Biehl, A., Leister, D., Rios, G., Koncz, C.*, et al.* (2004). Functional specialization amongst the Arabidopsis Toc159 family of chloroplast protein import receptors. Plant Cell *16*, 2059-2077.
Lamberti, G., Drurey, C., Soll, J., and Schwenkert, S. (2011a). The phosphorylation state of chloroplast transit peptides regulates preprotein import. Plant Signal Behav *6*, 1918-1920.

Lamberti, G., Drurey, C., Soll, J., and Schwenkert, S. (2011b). The phosphorylation state of chloroplast transit peptides regulates preprotein import. Plant Signaling & Behavior *6*, 1918-1920.
Lancelin, J., I, B., Arlaud, G., Blackledge, M., Gans, P., Stein, M., and Jaquot, J. (1994a). NMR structures of ferredoxin chloroplastic transit peptide from Chlamydomonas reinhardtii promoted by trifluoroethanol in aqueous solution. FEBS Letters *343*, 261-266.
Lancelin, J.M., Bally, I., Arlaud, G.J., Blackledge, M., Gans, P., Stein, M., and Jacquot, J.P. (1994b). NMR structures of ferredoxin chloroplastic transit peptide from Chlamydomonas reinhardtii promoted by trifluoroethanol in aqueous solution. FEBS letters *343*, 261-266.
Langner, U., Baudisch, B., and Klösgen, R.B. (2014). Organelle import of proteins with dual targeting properties into mitochondria and chloroplasts takes place by the general import pathways. Plant Signaling & Behavior *9*, e29301.
Lee, D.W., Jung, C., and Hwang, I. (2013). Cytosolic events involved in chloroplast protein targeting. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1833*, 245-252.
Lee, D.W., Lee, S., Lee, G.J., Lee, K.H., Kim, S., Cheong, G.W., and Hwang, I. (2006). Functional characterization of sequence motifs in the transit peptide of Arabidopsis small subunit of rubisco. Plant Physiol *140*, 466-483.
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.
Lee, D.W., Lee, S., Lee, G.J., Lee, K.H., Kim, S., Cheong, G.W., and Hwang, I. (2006). Functional characterization of sequence motifs in the transit peptide of Arabidopsis small subunit of rubisco. . Plant Physiology *140*, 466-483.
Lee, D.W., Lee, S., Oh, Y.J., Hwang, I. (2009). Multiple sequence motifs in the rubisco small subunit transit peptide independently contribute to Toc159 dependent import of proteins into chloroplasts. Plant Physiology *151*, 129-141.
Lee, J., Kim, D.H., and Hwang, I. (2014). Specific targeting of proteins to outer envelope membranes of endosymbiotic organelles, chloroplasts, and mitochondria. Frontiers in Plant Science *5*, 173.
Lee, J., Lee, H., Kim, J., Lee, S., Kim, D.H., Kim, S., and Hwang, I. (2011). Both the Hydrophobicity and a Positively Charged Region Flanking the C-Terminal Region of the Transmembrane Domain of Signal-Anchored Proteins Play Critical Roles in Determining Their Targeting Specificity to the Endoplasmic Reticulum or Endosymbiotic Organelles in Arabidopsis Cells. The Plant Cell *23*, 1588-1607.
Lee, J., Wang, F., and Schnell, D.J. (2009b). Toc receptor dimerization participates in the initiation of membrane translocation during protein import into chloroplasts. The Journal of biological chemistry *284*, 31130-31141.
Lee, K.H., Kim, D.H., Lee, S.W., Kim, Z.H., and Hwang, I. (2002). In vivo import experiments in protoplasts reveal the importance of the overall context but not specific amino acid residues of the transit peptide during import into chloroplasts. . Mol Cells *14*, 446-458.
Lee, S., Lee, D.W., Lee, Y., Mayer, U., Stierhof, Y.D., Jurgens, G., and Hwang, I. (2009c). Heat shock protein cognate 70-4 and an E3 ubiquitin ligase, CHIP, mediate

plastid-destined precursor degradation through the ubiquitin-26S proteasome system in Arabidopsis. Plant Cell *21*, 3984-4001.
Lee, U., Rioflorido, I., Hong, S.W., Larkindale, J., Waters, E.R., and Vierling, E. (2007). The Arabidopsis ClpB/Hsp100 family of proteins: chaperones for stress and chloroplast development. Plant J *49*, 115-127.
Lee, Y.J., Kim, D.H., Kim, Y.W., and Hwang, I. (2001). Identification of a signal that distinguishes between the chloroplast outer envelope membrane and the endomembrane system in vivo. Plant Cell *13*, 2175-2190.
Li, H., MM, K., PH, S., YH, Y., and CD, H. (2007). Toc GTPases. J Biomed Sci *14*, 505-508.
Li, H.M., and Chiu, C.C. (2010). Protein Transport into Chloroplasts. Annu Rev Plant Biol *61*, 157-180.
Li, Q.B., Anderson, J.V., and Guy, C.L. (1994). A cDNA clone encoding a spinach 70-kilodalton heat-shock cognate. Plant Physiol *105*, 457-458.
Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics *22*, 1658-1659.
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.
Lister, R., Chew, O., Rudhe, C., Lee, M.-N., and Whelan, J. (2001). Arabidopsis thaliana ferrochelatase-I and -II are not imported into Arabidopsis mitochondria. FEBS Letters *506*, 291-295.
Liu, L., McNeilage, R., Lan-xin, S., and Theg, S.M. (2014). ATP requirement for chloroplast protein import is set by Km for ATP hydrolysis of stromal HSP70 in Physcomitrella patens. The Plant Cell *26*, 1246-1255.
Lu, Y. (2016). Identification and Roles of Photosystem II Assembly, Stability, and Repair Factors in Arabidopsis. Frontiers in Plant Science *7*, 168.
Lumme, C., Altan-Martin, H., Dastvan, R., Sommer, Maik S., Oreb, M., Schuetz, D., Hellenkamp, B., Mirus, O., Kretschmer, J., Lyubenova, S*., et al.* (2014). Nucleotides and Substrates Trigger the Dynamics of the Toc34 GTPase Homodimer Involved in Chloroplast Preprotein Translocation. Structure *22*, 526-538.
Ma, Y., Kouranov, A., LaSala, S.E., and Schnell, D. (1996). Two components of the chloroplast import apparatus, IAP86 and IAP75, interact with the transit sequence during the recognition and translocation of precursor proteins at the outer envelope. J Cell Biol *134*, 315-327.
Marshall, J.S., DeRocher, A.E., Keegstra, K., and Vierling, E. (1990). Identification of heat shock protein hsp70 homologues in chloroplasts. Proc Natl Acad Sci U S A *87*, 374-378.
Martin, T., Sharma, R., Sippel, C., Waegemann, K., Soll, J., and Vothknecht, U.C. (2006). A protein kinase family in Arabidopsis phosphorylates chloroplast precursor proteins. J Biol Chem *281*, 40216-40223.
Martin, W., Brinkmann, H., Savonna, C., and Cerff, R. (1993). Evidence for a chimeric nature of nuclear genomes: eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. Proc Natl Acad Sci U S A *90*, 8692-8696.
Martin, W., and Kowallik, K.V. (1999). Annotated English translation of Mereschkowsky's 1905 paper 'Uber Natur und Ursprung der Chromatophoren im Pflanzenreiche. Eur J Phycol *34*, 287-295.

Mathur, J., Barton, K.A., and Schattat, M.H. (2013). Fluorescent protein flow within stromules. Plant Cell *8*, 2771-2772.
May, T., and Soll, J. (2000a). 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. Plant Cell *12*, 53-64.
May, T., and Soll, J. (2000b). 14-3-3 Proteins Form a Guidance Complex with Chloroplast Precursor Proteins in Plants. The Plant Cell *12*, 53-64.
McFadden, G.I. (2014). Origin and Evolution of Plastids and Photosynthesis in Eukaryotes. Cold Spring Harbor Perspectives in Biology *6*.
McFadden, G.I., and van Dooren, G.G. (2004). Evolution: red algal genome affirms a common origin of all plastids. Curr Biol *14*, R514-516.
Mitschke, J., Fuss, J., Blum, T., Höglund, A., Reski, R., Kohlbacher, O., and Rensing, S.A. (2009). Prediction of dual protein targeting to plant organelles. New Phytologist *183*, 224-236.
Moberg, P., Stahl, A., Bhushan, S., Wright, S.J., Eriksson, A., Bruce, B.D., and Glaser, E. (2003). Characterization of a novel zinc metalloprotease involved in degrading targeting peptides in mitochondria and chloroplasts. Plant J *36*, 616-628.
Mueller, K., Salnikow, J., and Vater, J. (1983). Amino acid sequence of the small subunit of D-ribulose bisphosphate carboxylase/oxygenase from Nicotiana tabacum. Biochim Biophys Acta *742*, 78-83.
Mullet, J. (1988). Chloroplast development and gene expression. Annu Rev Plant Physiol Plant Mol Biol *39*, 475-502.
Murakami, H., Pain, D., and Blobel, G. (1988). 70-kD heat shock-related protein is one of at least two distinct cytosolic factors stimulating protein import into mitochondria. J Cell Biol *107*, 2051-2057.
Murcha, M.W., Kmiec, B., Kubiszewski-Jakubiak, S., Teixeira, P.F., Glaser, E., and Whelan, J. (2014). Protein import into plant mitochondria: signals, machinery, processing, and regulation. Journal of Experimental Botany *65*, 6301-6335.
Nakrieko, K., Mould, R., and Smith, A. (2004). Fidelity of targeting to chloroplasts is not affected by removal of the phosphorylation site of the transit peptide. European Journal of Biochemistry *271*, 509-516.
Nanjo, Y., Oka, H., Ikarashi, N., Kaneko, K., Kitajima, A., Mitsui, T., Munoz, F.J., Rodriguez-Lopez, M., Baroja-Fernandez, E., and Pozueta-Romero, J. (2006). Rice plastidial N-glycosylated nucleotide pyrophosphatase/phosphodiesterase is transported from the ER-golgi to the chloroplast through the secretory pathway. Plant Cell *18*, 2582-2592.
Natesan, S.K.A., Sullivan, J.A., and Gray, J.C. (2005). Stromules: a characteristic cell-specific feature of plastid morphology. Journal of Experimental Botany *56*, 787-797.
Nelson, B.K., Cai, X., and Nebenfuhr, A. (2007). A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. Plant J *51*, 1126-1136.
Nickel, C., Soll, J., and Schwenkert, S. (2015). Phosphomimicking within the transit peptide of pHCF136 leads to reduced photosystem II accumulation in vivo. FEBS Letters *589*, 1301-1307.
Nielsen, E., Akita, M., Davila-Aponte, J., and Keegstra, K. (1997). Stable association of chloroplastic precursors with protein translocation complexes that contain proteins from

both envelope membranes and a stromal Hsp100 molecular chaperone. EMBO J *16*, 935-946.
Oblong, J.E., and Lamppa, G.K. (1992). Identification of two structurally related proteins involved in proteolytic processing of precursors targeted to the chloroplast. EMBO J *11*, 4401-4409.
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., Tews, I., Schnell, D.J., and Schleiff, E. (2011). Substrate binding disrupts dimerization and induces nucleotide exchange of the chloroplast GTPase Toc33. The Biochemical journal *436*, 313-319.
Pai, E.F., Kabsch, W., Krengel, U., Holmes, K.C., John, J., and Wittinghofer, A. (1989). Structure of the guanine-nucleotide-binding domain of the Ha-ras oncogene product p21 in the triphosphate conformation. Nature *341*, 209-214.
Paila, Y.D., Richardson, L.G.L., Inoue, H., Parks, E.S., McMahon, J., Inoue, K., and Schnell, D.J. (2016). Multi-functional roles for the polypeptide transport associated domains of Toc75 in chloroplast protein import. eLife *5*, e12631.
Pall, D., and Chakraabarti, P. (1999). Cis peptide bonds in proteins: residues involved, their conformation, interaction and locations. J Mol Biol *294*, 271-288.
Peeters, N., and Small, I. (2001). Dual targeting to mitochondria and chloroplasts. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1541*, 54-63.
Perrodou, E., Chica, C., Poch, O., Gibson, T., and Thompson, J. (2008). A new protein linear motif benchmark for multiple sequence alignment software. BMC Bioinformatics *9*.
Perry, S.E., and Keegstra, K. (1994). Envelope membrane proteins that interact with chloroplastic precursor proteins. Plant Cell *6*, 93-105.
Pesaresi, P., Masiero, S., Eubel, H., Braun, H.P., Bhushan, S., Glaser, E., Salamini, F., and Leister, D. (2006). Nuclear photosynthetic gene expression is synergistically modulated by rates of protein synthesis in chloroplasts and mitochondria. Plant Cell *18*, 970-991.
Pilon, M., de Boer, A.D., Knols, S.L., Koppelman, M.H., van der Graaf, R.M., de Kruijff, B., and Weisbeek, P.J. (1990). Expression in Escherichia coli and purification of a translocation-competent precursor of the chloroplast protein ferredoxin. J Biol Chem *265*, 3358-3361.
Pilon, M., de Kruijff, B., and Weisbeek, P. (1992a). New insights into the import mechanism of the ferredoxin precursor into chloroplasts. J Biol Chem *267*, 2548-2556.
Pilon, M., de Kruijff, B., and Weisbeek, P.J. (1992b). New insights into the import mechanism of the ferredoxin precursor into chloroplasts. J Biol Chem *267*, 2548-2556.
Pilon, M., Weisbeek, P., and de Kruijff, B. (1992c). Kinetic analysis of translocation into isolated chloroplasts of the purified ferredoxin precursor. FEBS Letters *302*.
Pilon, M., Wienk, H., Sips, W., de Swaaf, M., Talboom, I., van 't Hof, R., de Korte-Kool, G., Demel, R., Weisbeek, P., and de Kruijff, B. (1995a). Functional domains of the ferredoxin transit sequence involved in chloroplast import. J Biol Chem *270*, 3882-3893.
Pilon, M., Wienk, H., Sips, W., de Swaaf, M., Talboom, I., van't Hof, R., de Korte-Kool, G., Demel, R., Weisbeek, P., and de Kruijff, B. (1995b). Functional domains of the

ferredoxin transit sequence involved in chloroplast import. Journal of Biological Chemistry *270*, 3882-3893.
Pinnaduwage, P., and Bruce, B.D. (1996). In vitro interaction between a chloroplast transit peptide and chloroplast outer envelope lipids is sequence-specific and lipid class-dependent. The Journal of biological chemistry *271*, 32907-32915.
Plumley, F.G., and Schmidt, G.W. (1989). Nitrogen-dependent regulation of photosynthetic gene expression. Proc Natl Acad Sci U S A *86*, 2678-2682.
Primavesi, L.F., Wu, H., Mudd, E.A., Day, A., and Jones, H.D. (2007). 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 Research *17*, 529-543.
Qbadou, S., Becker, T., Mirus, O., Tews, I., Soll, J., and Schleiff, E. (2006). The molecular chaperone Hsp90 delivers precursor proteins to the chloroplast import receptor Toc64. EMBO J *25*, 1836-1847.
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.
Rapaport, D. (2003). Finding the right organelle: Targeting signals in mitochondrial outer-membrane proteins. EMBO reports *4*, 948-952.
Rapoport, T.A. (2007). Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. Nature *450*, 663-669.
Ratnayake, R.M.U., Inoue, H., Nonami, H., and Akita, M. (2008). Alternative Processing of Arabidopsis Hsp70 Precursors during Protein Import into Chloroplasts. Biosci Biotech Bioch *72*, 2926-2935.
Reddick, L.E. (2010). Dynamics of the Toc GTPases: Modulation by Nucleotides and Transit Peptides Reveal a Mechanism for Chloroplast Protein Import. Doctoral Dissertation, Vol Ph.D. (Knoxville: University of Tennessee).
Reddick, L.E., Chotewutmontri, P., Crenshaw, W., Dave, A., Vaughn, M., and Bruce, B.D. (2008a). Chapter 16 Nanoscale Characterization of the Dynamics of the Chloroplast Toc Translocon. In Methods in Cell Biology, P.J. Bhanu, ed. (Academic Press), pp. 365-398.
Reddick, L.E., Chotewutmontri, P., Crenshaw, W., Dave, A., Vaughn, M., and Bruce, B.D. (2008b). Nano-scale characterization of the dynamics of the chloroplast Toc translocon. Methods Cell Biol *90*, 365-398.
Reddick, L.E., Vaughn, M.D., Wright, S.J., Campbell, I.M., and Bruce, B.D. (2007a). In vitro comparative kinetic analysis of the chloroplast Toc GTPases. The Journal of biological chemistry *282*, 11410-11426.
Reddick, L.E., Vaughn, M.D., Wright, S.J., Campbell, I.M., and Bruce, B.D. (2007b). In Vitro Comparative Kinetic Analysis of the Chloroplast Toc GTPases. Journal of Biological Chemistry *282*, 11410-11426.
Rensink, W.A., Pilon, M., and Weisbeek, P. (1998). Domains of a transit sequence required for in vivo import in Arabidopsis chloroplasts. Plant Physiology *118*, 691-699.
Rensink, W.A., Schnell, D.J., and Weisbeek, P.J. (2000). The transit sequence of ferredoxin contains different domains for translocation across the outer and inner membrane of the chloroplast envelope. J Biol Chem *275*, 10265-10271.

Reumann, S., and Keegstra, K. (1999). The endosymbiotic origin of the protein import machinery of chloroplastic envelope membranes. Trends Plant Sci *4*, 302-307.
Rial, D.V., Arakaki, A.K., and Ceccarelli, E.A. (2000). Interaction of the targeting sequence of chloroplast precursors with Hsp70 molecular chaperones. Eur J Biochem *267*, 6239-6248.
Richardson, L.G.L., Paila, Y.D., Siman, S., Yi, C., Smith, M.D., and Schnell, D. (2014). Targeting and Assembly of Components of the TOC Protein Import Complex at the Chloroplast Outer Envelope Membrane. Frontiers in Plant Science *5*.
Richter, S., and Lamppa, G. (2002a). Determinants for removal and degradation of transit peptides of chloroplast precursor proteins. J Biol Chem *277*, 43888-43894.
Richter, S., and Lamppa, G.K. (1998). A chloroplast processing enzyme functions as the general stromal processing peptidase. Proceedings of the National Academy of Sciences of the United States of America *95*, 7463-7468.
Richter, S., and Lamppa, G.K. (1999). Stromal processing peptidase binds transit peptides and initiates their ATP-dependent turnover in chloroplasts. Journal of Cell Biology *147*, 33-43.
Richter, S., and Lamppa, G.K. (2002b). Determinants for removal and degradation of transit peptides of chloroplast precursor proteins. Journal of Biological Chemistry *277*, 43888-43894.
Richter, S., and Lamppa, G.K. (2003). Structural properties of the chloroplast stromal processing peptidase required for its function in transit peptide removal. Journal of Biological Chemistry *278*, 39497-39502.
Richter, S., Zhong, R., and Lamppa, G. (2005). Function of the stromal processing peptidase in the chloroplast import pathway. Physiol Plantarum *123*, 362-368.
Row, P.E., and Gray, J.C. (2001a). Chloroplast precursor proteins compete to form early import intermediates in isolated pea chloroplasts. J Exp Bot *52*, 47-56.
Row, P.E., and Gray, J.C. (2001b). The effect of amino acid-modifying reagents on chloroplast protein import and the formation of early import intermediates. J Exp Bot *52*, 57-66.
Rudhe, C., Chew, O., Whelan, J., and Glaser, E. (2002). A novel in vitro system for simultaneous import of precursor proteins into mitochondria and chloroplasts. The Plant Journal *30*, 213-220.
Scheffzek, K., Ahmadian, M.R., Kabsch, W., Wiesmuller, L., Lautwein, A., Schmitz, F., and Wittinghofer, A. (1997). The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. Science *277*, 333-338.
Schleiff, E., Soll, J., Kuchler, M., Kuhlbrandt, W., and Harrer, R. (2003a). Characterization of the translocon of the outer envelope of chloroplasts. J Cell Biol *160*, 541-551.
Schleiff, E., Soll, J., Kuechler, M., Kuelbrandt, W., and Harrer, R. (2003b). Characterization of the translocon of the outer envelope of chloroplasts. J Cell Biol *160*, 541-551.
Schleiff, E., Soll, J., Sveshnikova, N., Tien, R., Wright, S., Dabney-Smith, C., Subramanian, C., and Bruce, B.D. (2002a). Structural and guanosine triphosphate/diphosphate requirements for transit peptide recognition by the cytosolic domain of the chloroplast outer envelope receptor, Toc34. Biochemistry *41*, 1934-1946.

Schleiff, E., Soll, J., Sveshnikova, N., Tien, R., Wright, S., Dabney-Smith, C., Subramanian, C., and Bruce, B.D. (2002b). Structural and guanosine triphosphate/diphosphate requirements for transit peptide recognition by the cytosolic domain of the chloroplast outer envelope receptor, Toc34. Biochemistry *41*, 1934-1946.
Schmidt GW, D.-T.A., Desruisseaux H, Blobel G, Chua NH. (1979). NH2-terminal amino acid sequences of precursor and mature forms of the ribulose-1,5-bisphosphate carboxylase small subunit from Chlamydomonas reinhardtii. Journal of Cell Biology *83*, 615-622.
Schnell, D., and Blobel, G. (1993a). Identification of intermediates in the pathway of protein import into chlroplasts and their localization to envelope contact sites. J Cell Biol *120*, 103-115.
Schnell, D., Blobel, G., and Pain, D. (1990). The chloroplast import receptor is an integral membrane protein of chloroplast envelope contact sites. J Cell Biol *111*, 1825-1838.
Schnell, D.J., and Blobel, G. (1993b). Identification of intermediates in the pathway of protein import into chloroplasts and their localization to envelope contact sites. J Cell Biol *120*, 103-115.
Schnell, D.J., Blobel, G., Keegstra, K., Kessler, F., Ko, K., and Soll, J. (1997). A consensus nomenclature for the protein-import components of the chloroplast envelope. Trends Cell Biol *7*, 303-304.
Schnell, D.J., Kessler, F., and Blobel, G. (1994). Isolation of components of the chloroplast protein import machinery. Science *266*, 1007-1012.
Schulz, C., Schendzielorz, A., and Rehling, P. (2015). Unlocking the presequence import pathway. Trends in Cell Biology.
Scott, S.V., and Theg, S.M. (1996). A new chloroplast protein import intermediate reveals distinct translocation machineries in the two envelope membranes: energetics and mechanistic implications. The Journal of Cell Biology *132*, 63-75.
Seedorf, M., and Soll, J. (1995). Copper chloride, an inhibitor of protein import into chloroplasts. FEBS Lett *367*, 19-22.
Shen, G., Adam, Z., and Zhang, H. (2007). The E3 ligase AtCHIP ubiquitylates FtsH1, a component of the chloroplast FtsH protease, and affects protein degradation in chloroplasts. Plant J *52*, 309-321.
Shi, L.-X., and Theg, S.M. (2010a). A stromal heat shock protein 70 system functions in protein import into chloroplasts in the moss Physcomitrella patens. The Plant Cell *22*, 205-220.
Shi, L.-X., and Theg, S.M. (2013). Energetic cost of protein import across the envelope membranes of chloroplasts. Proceedings of the National Academy of Sciences *110*, 930-935.
Shi, L.X., and Theg, S.M. (2010b). A Stromal Heat Shock Protein 70 System Functions in Protein Import into Chloroplasts in the Moss Physcomitrella patens. The Plant cell *22*, 205-220.
Shi, L.X., and Theg, S.M. (2011). The motors of protein import into chloroplasts. Plant Signal Behav *6*, 1397-1401.
Silva-Filho, M.C. (2003). One ticket for multiple destinations: dual targeting of proteins to distinct subcellular locations. Current Opinion in Plant Biology *6*, 589-595.

Sirajuddin, M., Farkasovsky, M., Zent, E., and Wittinghofer, A. (2009). GTP-induced conformational changes in septins and implications for function. Proceedings of the National Academy of Sciences of the United States of America *106*, 16592-16597.
Smeekens, S., van Binsbergen, J., and Weisbeck, P. (1985). The plant ferredoxin precursor: nucleotide sequence of a full length cDNA clone Nucleic Acids Res *13*, 3179-3194.
Smith, M.D. (2006). Protein import into chloroplasts: an ever-evolving story. Can J Bot *84*, 531-542.
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.
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.
Soll, J., and Schleiff, E. (2004). Protein import into chloroplasts. Nat Rev Mol Cell Biol *5*, 198-208.
Song, J., Burrage, K., Yuan, Z., and Huber, T. (2006). Prediction of cis/trans isomerization in proteins using PSI-BLAST profiles and secondary structure information. BMC Bioinformatics *7*.
Sprang, S. (1997). G proteins, effectors and GAPS: structure and mechanism. Current opinion in structural biology *7*, 849-856.
Stahl, A., Moberg, P., Ytterberg, J., Panfilov, O., Brockenhuus Von Lowenhielm, H., Nilsson, F., and Glaser, E. (2002). Isolation and identification of a novel mitochondrial metalloprotease (PreP) that degrades targeting presequences in plants. J Biol Chem *277*, 41931-41939.
Stengel, A., Benz, J.P., Buchanan, B.B., Soll, J., and Bolter, B. (2009). Preprotein import into chloroplasts via the Toc and Tic complexes is regulated by redox signals in Pisum sativum. Molecular plant *2*, 1181-1197.
Stengel, A., Benz, P., Balsera, M., Soll, J., and Bolter, B. (2008). TIC62 redox-regulated translocon composition and dynamics. J Biol Chem *283*, 6656-6667.
Su, P.H., and Li, H.M. (2008). Arabidopsis stromal 70-kD heat shock proteins are essential for plant development and important for thermotolerance of germinating seeds. Plant Physiol *146*, 1231-1241.
Su, P.H., and Li, H.M. (2010). Stromal Hsp70 Is Important for Protein Translocation into Pea and Arabidopsis Chloroplasts. Plant Cell *22*, 1516-1531.
Subramanian, C., Ivey, R.A.I., and Bruce, B.D. (2001). Technical advance: cytometric analysis of an epitope tagged transit peptide bound to the chloroplast translocation apparatus. The Plant Journal *25*, 349-363.
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.
Sun, Y.-J., Forouhar, F., Li, H.-m., Tu, S.-L., Yeh, Y.-H., Kao, S., Shr, H.-L., Chou, C.-C., Chen, C., and Hsiao, C.-D. (2002a). Crystal structure of pea Toc34, a novel GTPase of the chloroplast protein translocon. Nat Struct Biol *9*, 95-100.

Sun, Y.J., Forouhar, F., Li Hm, H.M., Tu, S.L., Yeh, Y.H., Kao, S., Shr, H.L., Chou, C.C., Chen, C., and Hsiao, C.D. (2002b). Crystal structure of pea Toc34, a novel GTPase of the chloroplast protein translocon. Nat Struct Biol *9*, 95-100.
Sung, D.Y., Vierling, E., and Guy, C.L. (2001). Comprehensive expression profile analysis of the Arabidopsis Hsp70 gene family. Plant Physiol *126*, 789-800.
Sveshnikova, N., Soll, J., and Schleiff, E. (2000). Toc34 is a pre-protein receptor regulated by GTP and phosphorylation. Proceedings of the National Academy of Science USA *97*, 4973-4978.
Teng, Y.S., Chan, P.T., and Li, H.M. (2012). Differential age-dependent import regulation by signal peptides. PLoS Biol *10*, e1001416.
Teng, Y.S., Su, Y.S., Chen, L.J., Lee, Y.J., Hwang, I., and Li, H.M. (2006). Tic21 is an essential translocon component for protein translocation across the chloroplast inner envelope membrane. Plant Cell *18*, 2247-2257.
Theg, S.M., Bauerle, C., Olsen, L.J., Selman, B.R., and Keegstra, K. (1989a). Internal ATP is the only energy requirement for the translocation of precursor proteins across chloroplastic membranes. Journal of Biological Chemistry *264*, 6730-6736.
Theg, S.M., Bauerle, C., Selman, B.R., and Keegstra, K. (1989b). Internal ATP is the only energy requirement for the translocation of precursor proteins across chloroplastic membranes. J Biol Chem *264*, 6730-6736.
Tonkin, C.J., Foth, B.J., Ralph, S.A., Struck, N., Cowman, A.F., and McFadden, G.I. (2008). Evolution of malaria parasite plastid targeting sequences. Proceedings of the National Academy of Sciences *105*, 4781-4785.
Tranel, P.J., and Keegstra, K. (1996). A novel, bipartite transit peptide targets OEP75 to the outer membrane of the chloroplastic envelope. Plant Cell *8*, 2093-2104.
Tsai, L.Y., Tu, S.L., and Li, H.M. (1999). Insertion of atToc34 into the chloroplastic outer membrane is assisted by at least two proteinaceous components in the import system. J Biol Chem *274*, 18735-18740.
van 't Hof, R., and de Kruijff, B. (1995a). Characterization of the import process of a transit peptide into chloroplasts. J Biol Chem *270*, 22368-22373.
van 't Hof, R., and de Kruijff, B. (1995b). Transit sequence-dependent binding of the chloroplast precursor protein ferredoxin to lipid vesicles and its implications for membrane stability. FEBS Lett *361*, 35-40.
van 't Hof, R., Demel, R.A., Keegstra, K., and de Kruijff, B. (1991). Lipid-peptide interactions between fragments of the transit peptide of ribulose-1,5-bisphosphate carboxylase/oxygenase and chloroplast membrane lipids. FEBS Lett *291*, 350-354.
van 't Hof, R., van Klompenburg, W., Pilon, M., Kozubek, A., de Korte-Kool, G., Demel, R.A., Weisbeek, P.J., and de Kruijff, B. (1993). The transit sequence mediates the specific interaction of the precursor of ferredoxin with chloroplast envelope membrane lipids. J Biol Chem *268*, 4037-4042.
van den Wijngaard, P.W., Dabney-Smith, C., Bruce, B.D., and Vredenberg, W.J. (1999). The mechanism of inactivation of a 50-pS envelope anion channel during chloroplast protein import. Biophys J *77*, 3156-3162.
van den Wijngaard, P.W.J., and Vredenberg, W.J. (1999). The envelope anion channel involved in chloroplast protein import is associated with Tic110. Journal of Biological Chemistry *274*, 25201-25204.

VanderVere, P.S., Bennett, T.M., Oblong, J.E., and Lamppa, G.K. (1995). A chloroplast processing enzyme involved in precursor maturation shares a zinc-binding motif with a recently recognized family of metalloendopeptidases. Proc Natl Acad Sci U S A *92*, 7177-7181.
Villarejo, A., Buren, S., Larsson, S., Dejardin, A., Monne, M., Rudhe, C., Karlsson, J., Jansson, S., Lerouge, P., Rolland, N.*, et al.* (2005). Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. Nature cell biology *7*, 1224-1231.
Voigt, A., Jakob, M., Klosgen, R.B., and Gutensohn, M. (2005). At least two Toc34 protein import receptors with different specificities are also present in spinach chloroplasts. FEBS Lett *579*, 1343-1349.
Vojta, A., Alavi, M., Becker, T., Hormann, F., Kuchler, M., Soll, J., Thomson, R., and Schleiff, E. (2004). The protein translocon of the plastid envelopes. J Biol Chem *279*, 21401-21405.
von Heijne, G., and Nishikawa, K. (1991). Chloroplast transit peptides. The perfect random coil? *FEBS Lett*, 1-3.
von Heijne, G., Steppuhn, J., and Herrmann, R.G. (1989a). Domain structure of mitochondrial and chloroplast targeting peptides. Eur J Biochem *180*, 535-545.
von Heijne, G., Steppuhn, J., and Herrmann, R.G. (1989b). Domain structure of mitochondrial and chloroplast targeting peptides. European Journal of Biochemistry *180*, 535-545.
Vorst, O., van Dam, F., Oosterhoff-Teertstra, R., Smeekens, S., and Weisbeek, P. (1990). Tissue-specific expression directed by an Arabidopsis thaliana pre-ferredoxin promoter in transgenic tobacco plants. Plant Mol Biol *14*, 491-499.
Vothknecht, U.C., and Soll, J. (2005). Chloroplast membrane transport: Interplay of prokaryotic and eukaryotic traits. Gene *354*, 99-109.
Waegemann, K., Paulsen, H., and Soll, J. (1990). Translocation of Proteins into Isolated-Chloroplasts Requires Cytosolic Factors to Obtain Import Competence. Febs Letters *261*, 89-92.
Waegemann, K., and Soll, J. (1996). Phosphorylation of the transit sequence of chloroplast precursor proteins. J Biol Chem *271*, 6545-6554.
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.
Walter, P., and Lingappa, V.R. (1986). Mechanism of protein translocation across the endoplasmic reticulum membrane. Annual review of cell biology *2*, 499-516.
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.
Waterhouse, A., Proctor, J., Martin, D., Clamp, M., and Barton, G. (2009). Jalview Version 2: a mutliple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189-1191.
Weiss, M.S., Jabs, A., and Hilgenfeld, R. (1998). Peptide bonds revisited. Nat Struct Biol *5*, 676.
Westphal, S., Soll, J., and Vothknecht, U.C. (2003). Evolution of Chloroplast Vesicle Transport. Plant and Cell Physiology *44*, 217-222.

Wienk, H., Czisch, M., deKruijff, B. (1999). The structural flexibility of the preferredoxin transit peptide. FEBS Letters *453*, 318-326.
Wienk, H.L., Czisch, M., and de Kruijff, B. (1999). The structural flexibility of the preferredoxin transit peptide. FEBS letters *453*, 318-326.
Wienk, H.L., Wechselberger, R.W., Czisch, M., and de Kruijff, B. (2000). Structure, dynamics, and insertion of a chloroplast targeting peptide in mixed micelles. Biochemistry *39*, 8219-8227.
Yan, X., Khan, S., Hase, T., Emes, M.J., and Bowsher, C.G. (2006). Differential uptake of photosynthetic and non-photosynthetic proteins by pea root plastids. FEBS Lett *580*, 6509-6512.
Yeh, Y.H., Kesavulu, M.M., Li, H.M., Wu, S.Z., Sun, Y.J., Konozy, E.H., and Hsiao, C.D. (2007). Dimerization is important for the GTPase activity of chloroplast translocon components atToc33 and psToc159. J Biol Chem *282*, 13845-13853.
Yi-Shan, T., Po-Ting, C., and Hsou-min, L. (2012). Differential Age-Dependent Import Regulation by Signal Peptides. PLoS Biology *10*, 1-14.
Yoon, H.S., Hackett, J.D., Ciniglia, C., Pinto, G., and Bhattacharya, D. (2004). A molecular timeline for the origin of photosynthetic eukaryotes. Molecular biology and evolution *21*, 809-818.
Young, M.E., Keegstra, K., and Froehlich, J.E. (1999). GTP promotes the formation of early-import intermediates but is not required during the translocation step of protein import into chloroplasts. Plant Physiol *121*, 237-244.
Zhang, X.P., and Glaser, E. (2002). Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone. Trends in plant science *7*, 14-21.
Zhang, Y.e.a. (2011). A highly efficient rice green tissue protoplast system for transient gene expression and studying light/chloroplast-related processes. Plant Methods *7*, 30-43.
Zhong, R., Wan, J., Jin, R., and Lamppa, G. (2003). A pea antisense gene for the chloroplast stromal processing peptidase yields seedling lethals in Arabidopsis: survivors show defective GFP import in vivo. Plant J *34*, 802-812.
Zhou, Y.H., Yu, J.Q., Mao, W.H., Huang, L.F., Song, X.S., and Nogues, S. (2006). Genotypic variation of Rubisco expression, photosynthetic electron flow and antioxidant metabolism in the chloroplasts of chill-exposed cucumber plants. Plant & cell physiology *47*, 192-199.
Zimmermann, R., Sagstetter, M., Lewis, M.J., and Pelham, H.R. (1988). Seventy-kilodalton heat shock proteins and an additional component from reticulocyte lysate stimulate import of M13 procoat protein into microsomes. EMBO J *7*, 2875-2880.
Zybailov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q., and van Wijk, K.J. (2008). Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. PLoS One *3*, e1994.

APPENDIX

Appendix 1

Generated Construct	Primer Name	Primer Sequence (5'-3')	Template
psToc34 Q71A	Q71A Up / RC Q71A Lo	**UP:** G GTG GTT TCA ATT AGT CCC TTT GCG TCT GAA GGG CC **LO:** GGC CCT TCA GAC GCA AAG GGA CTA ATT GAA ACC	pET21D psToc34 △TM2
psToc34 Q71E	Q71E Up/ RC Q7E Lo	**UP:** G GTG GTT TCA ATT AGT CCC TTT GAG TCT GAA GGG CC **LO:** GGC CCT TCA GAC TCA AAG GGA CTA ATT GAA ACC ACC	pET21D psToc34 △TM2
psToc34 E73A	E73A Up/ RC E73A Lo	**UP:** CC TTT CAG TCT GCA GGG CCA AGA CCT GTT ATG G **LO:** CCA TAA CAG GTC TTG GCC CTG CAG ACT GAA AGG	pET21D psToc34 △TM2
psToc34 E73D	E73D Up/ RC E73D Lo	**UP:** CC TTT CAG TCT GAC GGG CCA AGA CCT GTT ATG G **LO:**CCA TAA CAG GTC TTG GCC CGT CAG ACT GAA AGG	pET21D psToc34 △TM2
psToc34 E73Q	E73Q Up/RC E73Q Lo	**UP:** CC TTT CAG TCT CAA GGG CCA AGA CCT GTT ATG G **LO:** CCA TAA CAG GTC TTG GCC CTT GAG ACT GAA AGG	pET21D psToc34 △TM2
psToc34Q71E/E73G/G74E	E73/G74E Up/ RC E73G/G74E Lo		pET21D psToc34 Q71E △TM2
psToc34 Q71E/E73S	E73S Up/RC E73S Lo	**UP:** CCC TTT CAG TCT TCA GGG CCA AGA CCT GTT ATG G **LO:** CCA TAA CAG GTC TTG GCC CTG AAG ACT GAA	pET21D psToc34 Q71E △TM2

Appendix 2

Generated Construct	Primer Name	Primer Sequence (5'-3')
DI 5S	FGLK-5S UP/FGLK-5S LO	**UP**: GCA CCT TCC TCT TCC TCT AGC TCA GCT GCC **LO**: GGC AGC TGA GCT AGA GGA AGA GGA AGG TGC
DII 5S	FPVSR-5S UP/FPVSR-5S LO	**UP**: GCT GCC TCA TCC TCT TCT TCA AGC AAG C **LO**:GCT TGC TTG AAG AAG AGG ATG AGG CAG C
DI-N 3S	FTG-3S UP/FTG-3S LO	**UP**: GCA CCT TCC TCT TCC CTT AAG TCA GCT GCC **LO:** GGC AGC TGA CTT AAG GGA AGA GGA AGG TGC
DI-C 3S	GLK-3S UP/GLK-3S LO	**UP:** GCA CCT TTC ACT TCC TCT AGC TCA GCT GCC **LO**:GGC AGC TGA GCT AGA GGA AGT GAA AGG TGC
DII-N 3S	FPV-3S UP/FPV-3S LO	**UP**:GCT GCC TCA TCC TCT TCT TCA AGG AAG C **LO**:GCT TCC TTG AAG AAG AGG ATG AGG CAG C
DII-C 3S	VSR-3S UP/VSR-3S LO	**UP**:GCT GCC TCA TTC CCT TCT TCA AGC AAG C **LO**:GCT TGC TTG AAG AAG GGA ATG AGG CAG C
DI-K30S	FTGLS-UP/FTGLS-LO	**UP**:GCA CCT TTC ACT GGC CTT AGC TCA GCT GCC **LO**:GGC AGC TGA GCT AAG GCC AGT GAA AGG TGC
DII-RK 2S	FPVSSS-RK-UP/FPVSSS-RK-LO	**UP**:C CCT GTT TCA AGC TCC CAA AAC C **LO:** GGT TTT GGG AGC TTG AAA CAG GG
S31/34A	S31-34A-UP/S31-34A-LO	**UP**:GGC CTT AAG GCT GCT GCC GCT TTC CC **LO**:GGG AAA GCG GCA GCA GCC TTA AGG CC
S31/34D	S31-34D-UP/S31-34D-LO	**UP**:GGC CTT AAG GAC GCT GCC GAC TTC CCT GTT TCA AGG AAG **LO**:CTT CCT TGA AAC AGG GAA GTC GGC AGC GTC CTT AAG GCC
T27D	FDGLK UP/FDGLK LO	**UP**:CCT TTC GAC GGC CTT AAG TCA GC **LO:** GCT TCC TGT CAA CAG GGA ATG AGG C
S38D	FPVDR UP/FPVDR LO	**UP**: GCC TCA TTC CCT GTT GAC AGG AAGC **LO**: GCT TCC TGT CAA CAG GGA ATG AGG C
R14A	R14A UP/R14A LO	**UP**: GCCACCGCCAGCAATGTTGCTCAAGC **LO**:GCTTGAGCAACATTGCTGGCGGTGGC
R54A	R54A UP/R54A LO	**UP**: AGCAACGGCGGAGCTGTGCAATGC **LO**:GCATTGCACAGCTCCGCCGTTGCT
DIII-4S (FDF)	FDF -FGLK-5S UP/FGLK-5S LO	**UP**: CCAGGCACTGTCTTCTTCTTCTGCAGGTTCTCGC **LO**:GCGAGAACCTGCAGAAGAAGAAGACAGTGCCTGG
DI-6S (FDF)	FDF-PKQQPM-6S UP/PKQQPM 6S LO	**UP**: TCT CTG CTG TCT TCC TCC TCC TCT TCC GTG GC **LO**: GCC ACG GAA GAG GAG GAG GAA GAC AGC AGA GA
DI-N 3S (FDF)	FDF PKQ-3S UP/PKQ 3S LO	**UP**: GCT TCT CTG CTG TCT TCC TCC CAG CCG ATG GTG GCA TC **LO**: GAT GCC ACC ATC GGC TGG GAG GAA GAC AGC AGA GAA GC
DI-C 3S (FDF)	FDF QPM 3S UP/QPM 3S LO	**UP**: GCT TCT CTG CTG CCG AAG CAG TCC TCT TCC GTG GCA TC **LO**: GAT GCC ACG GAA GAG GAC TGC TTC GGC AGC AGA GAA GC

Appendix 3

Generated Construct	Primer Name	Primer Sequence (5'-3')
TP1	PEPT1-UP-NHE1/PEPT1-LO-NCO1	**UP**:GCT AGC ATG GCT GCT ACT TCT TCT ATT **LO**: CCA TGG A CAA AGG AAA AGA AGG AGA AGA
TP2	PEPT2-UP-NHE1/PEPT2-LO-NCO1	**UP**: GCT AGC ATG GCT ACT CAT TCT ATG ATT **LO**: CCA TGG A ATC CAA TTG AGA ATG CTT TTC
TP3	PEPT3-UP-NHE1/PEPT3-LO-NCO1	**UP**: GCT AGC ATG GCT TCT TTG TTG GGA ACT **LO**: CCA TGG A AGA ATT AAT TTC AGT AGC AAC
TP4	PEPT4-UP-NHE1/PEPT4-LO-NCO1	**UP**: GCT AGC ATG GCT ACT TCT TCT ACT CAA **LO**: CCA TGG A TTC ATT AAC ATC AAC AGC AGA
TP5	PEPT5-UP-NHE1/PEPT5-LO-NCO1	**UP**: GCT AGC ATG GCT GCT GGA TTG TCT ACT **LO**: CCA TGG A ATC CAT AGG AGA AGA TTG CTT
TP6	PEPT6-UP-NHE1/PEPT6-LO-NCO1	**UP**: GCT AGC ATG ATT ATG GCT TCT TCT GCT **LO**: CCA TGG A TCC TCC AGT AGA AGA AGT AGT
TP7	PEPT7-UP-NHE1/PEPT7-LO-NCO1	**UP**: GCT AGC ATG GCT GCT GCT GTT TCT ACT **LO**: CCA TGG A TCC ATC AGT TTG CTT ATC TTC

VITA

Kristen Holbrook was born in Johnson City, Tennessee on February 5, 1985. She was raised in Kingsport, Tennessee and graduated from Sullivan South High School in 2003. She obtained her Honors Cum Laude Bachelor's Degree (Biology/concentration: Biochemistry) from the University of Tennessee in 2007. She received her Master's Degree (Biology/concentration: Biochemistry) in 2012 from the laboratory of Dr. Rose Goodchild. She joined Dr. Barry Bruce's lab in Winter 2013 and began studying chloroplast protein import. She plans to continue to study phototrophic organisms as a postdoctoral researcher.