2012

Toward a functional characterization of the Acidic-domain of the chloroplast protein import receptor Toc159

Joseph K. Weston
Wilfrid Laurier University, west8730@gmail.com

Toward a Functional Characterization of the Acidic-domain of the Chloroplast Protein Import Receptor Toc159

by

J. Kyle Weston

A thesis
presented to Wilfrid Laurier University
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Integrative Biology

Waterloo, Ontario, Canada, 2012

I declare that I am the sole author of this thesis.

I understand that this work may be made available to the public

Abstract

Chloroplasts are members of a diverse class of organelles called plastids that differentiate plants from other eukaryotes, and are the site of a number of essential biochemical pathways including photosynthesis. Nuclear-encoded pre-proteins, which account for ~95% of chloroplast proteins, are post-translationally imported into plastids across the double envelope membrane. In the model plant *Arabidopsis thaliana*, the majority of pre-proteins are imported via the Toc (**t**ranslocon at **o**uter envelope membrane of **c**hloroplasts) and Tic (**t**ranslocon at **i**nner envelope membrane of **c**hloroplasts) complexes, the key components of which have been identified. The Toc159 homologues, atToc159 and atToc132/120, have been shown to form structurally and functionally distinct Toc complexes and have been proposed to serve as the primary pre-protein receptors, recognizing and interacting with the N-terminal extensions of pre-proteins, called transit peptides. The tripartite structure of atToc159 includes a membrane anchor (M-) domain, a GTPase (G-) domain, and an acidic (A-) domain that is currently functionally uncharacterized. The A-domain is a large (733 a.a.), intrinsically unstructured protein domain. The sequence identity among Toc159 homologue A-domains is considerably lower than the G- and M-domains. The A-domain has been shown to play a role in the pre-protein specificity exhibited by distinct Toc complexes. In the current study, the A-domain of atToc159 was shown to remain associated with the chloroplast envelope membrane when proteolytically cleaved by thrombin and to interact specifically with the G-domain of atToc159 *in vitro*. There was not a large change in secondary structure associated with the interaction, as observed by CD, but the interaction between the A- and G-domains of atToc159 was observed to inhibit the hydrolysis of GTP by the G-domain.

Acknowledgements

I would like to thank my supervisor, Dr. Matt Smith, for his patience and guidance and for putting up with my erratic behaviour. I would also like to thank Dr. Masoud Jelokhani-Niaraki for his assistance and guidance in the use of CD and the analysis of the data and Dr. Simon Chuong for his help and insight with this project.

I would like to thank Lynn Richardson for her previous work on the A-domain, Patrick Hoang for his friendship and helpful insight, Spence Macdonald for keeping things in the lab fun, Dr. Sid Dutta for his encyclopedia-like brain, as well as all past and present members of the Smith Lab that have contributed to my experience over the last three years. A special thanks goes to my family, especially my mother, sisters, and my brother, for keeping me sane.

Table of Contents

List of Figures

List of Abbreviations

atToc159A: protein corresponding to the Acidic-domain of the atToc159 homologue

atToc132A: protein corresponding to the Acidic-domain of the atToc132 homologue

atToc159G: protein corresponding to the GTPase-domain of the atToc159 homologue

atToc132G: protein corresponding to the GTPase-domain of the atToc132 homologue

AKR2A: ankyrin repeat protein 2A

CD: circular dichroism spectropolarimetry

ceQORH: chloroplast envelope quinone Oxidoreductase

F-actin: Human, non-muscle filamentous actin

GAP: GTPase activating proteins

GDF: GDP dissociation inhibitor displacement factor

GDI:GDP dissociation inhibitor

GTP: guanosine triphosphate

Hip/Hop domain: Hsc70-interacting protein/Hsc70/Hsp90-organizing protein domain

IEM: inner envelope membrane

IMAC: immobilized metal affinity chromatography

IMS: intermembrane space

IPTG: Isopropyl β-D-1-thiogalactopyranoside

IUP: intrinsically unstructured protein

IVT: in vitro translation product

MGDG: monogalactosyldiacyl glycerol

OEP: outer envelope protein

OEM: outer envelope membrane

*ppi*1(2,3): plastid protein import mutant 1 (2,3)

pSSU: precursor of the small subunit of Rubisco

pE1α – precursor of the E1α subunit of pyruvate dehydrogenase

TA: tail-anchored

THP: Tris(hydroxypropyl) phosphine

Tic: translocon at the inner membrane of chloroplasts

TMD: transmembrane domain

Toc: translocon at the outer membrane of chloroplasts

1. Introduction

1.1 Plastids

Plastids are a family of semi-autonomous organelles unique to plants that possess a number of structurally and functionally diverse constituents. All plastid types are derived from a common precursor, the proplastid, found largely in rapidly dividing cells like embryonic and meristematic tissues. As these cells differentiate towards their functional endpoints, the proplastids differentiate into functionally distinct plastid types. While all plastids share physical similarities, such as a double envelope membrane, they differ in their internal membrane systems and protein complements (Wise 2006).

Differentiated, functional plastid types include: etioplasts, a chloroplast-like plastid that lacks photosynthetic machinery in dark conditions, but can develop photosynthetic apparatus in the presence of light (Wise 2006); amyloplasts, a type of leucoplasts important for the storage of starch granules and graviperception by the movement of starch granules (Wise 2006); elaioplasts, the site of "prokaryotic pathway" lipid synthesis – e.g. triacylglycerols, monogalactosyldiacylglycerols, and sterol esters. The storage of these lipids in the form of oil droplets, and the development of the pollen coat in the final stages of maturation (Wise 2006); proteinoplasts, although their role in anabolism, catabolism or storage is not clearly defined, it is clear that these plastids contain an extensive protein complement and contain large inclusion bodies (Wise 2006); chromoplasts, carotenoid-storage plastids that develop from chloroplasts, generally found in brightly coloured tissue of flowers

and fruits; and gerontoplasts, found in senescing green tissue, these plastids also arise from chloroplasts (Wise 2006).

1.2 Chloroplast Structure and Function

While all members of the plastid family are crucial for plant development and function, the most well known and best studied plastid is the chloroplast. Chloroplasts are abundant in leafy, green tissues, often exposed to light, and are the site of photosynthesis (Wise 2006). While photosynthesis is an important and well described chloroplastic process, chloroplasts are also the site of nitrogen assimilation, sulphur metabolism, and the synthesis of a number of pigments, fatty acids, and amino acids (Wise 2006). Chloroplasts possess an intricate internal membrane system, that when coupled with the double envelope membrane shared by all plastids, gives rise to six separate sub-organellar compartments. The chloroplast envelope double-membrane provides four of these sub-compartments: the outer envelope membrane, intermembrane space, the inner envelope membrane, and the stroma (Smith & Schnell 2004). The highly developed thylakoid membrane system, comprised of stacks of thylakoids, called grana, connected by an intricate network of lamellae, essential to the chloroplast's photosynthetic reactions contributes the other two sub-compartments: the thylakoid membrane and the thylakoid lumen (Figure 1; Smith & Schnell 2004).

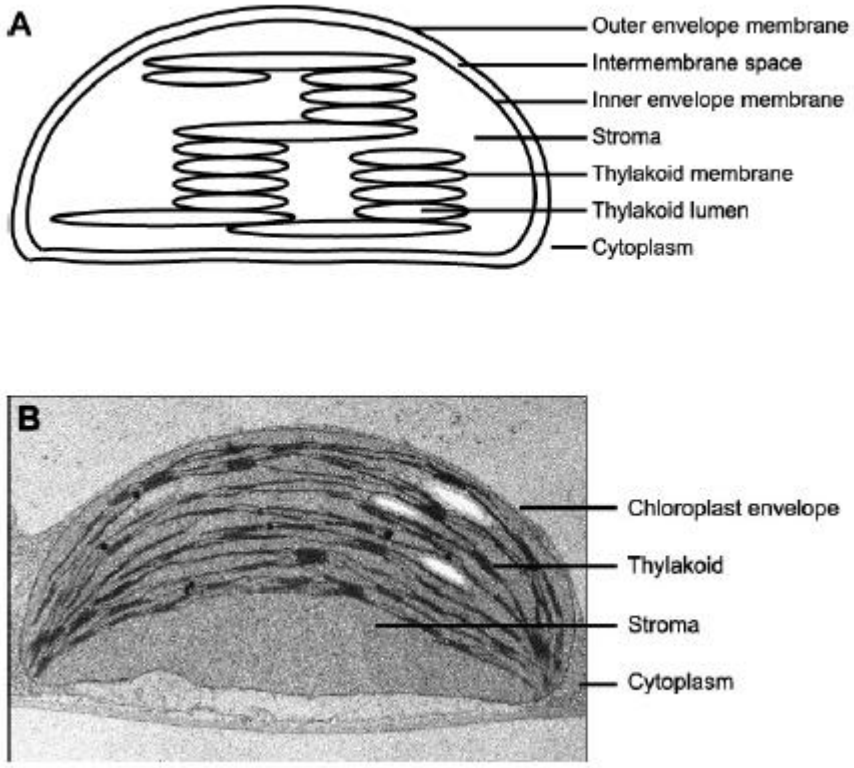


Figure 1. The general structure of chloroplasts. The diagram in (A) illustrates the six distinct sub-compartments of the chloroplast. The transmission electron micrograph of an *Arabidopsis thaliana* leaf cell depicts chloroplast shape and structure (Smith & Schnell, 2004).

1.3 Endosymbiotic Origin of Chloroplasts

It is accepted that plastids arose from an endosymbiotic event. The first such event involved the uptake of a cyanobacterium by a mitochondria-containing eukaryote (Dyall et al, 2004; Keeling, 2004). This cyanobacterium was retained within the cell, allowed to self replicate, and over time a significant gene transfer to the host nuclear genome occurred. A second, later endosymbiotic event, involving the uptake of this photosynthetic eukaryote by a non-photosynthetic eukaryote, led to the development of organisms with so-called secondary plastids with more complex membrane systems (Thomson and Whatley, 1980). The cyanobacterial plastid ancestor is most similar to chloroplasts, largely due to their photosynthetic capability. However, it is the ability of plants to maintain non-photosynthetic plastids in certain life stages and tissues that facilitated the evolution of land plants (Thomson and Whatley, 1980). It has been suggested that by reducing unnecessary energy used for the biogenesis and maintenance of photosynthetic machinery, plastids could be utilized in other capacities (Thomson and Whatley, 1980). The reduced genome of present-day chloroplasts codes for approximately 120 genes, most of which are involved in translation and photosynthesis (Wise, 2006). The chloroplast genome is significantly smaller than the genome of the unicellular cyanobacteria, *Synechocystis* sp, whose genome codes for over 3000 genes (Kaneko et al. 1996; Raven & Allen. 2003). The reduction of the endosymbiont's genome was a result of significant gene transfer from the plastid genome to the host nuclear genome (Martin & Herrmann 1998; Raven & Allen 2003). For the proteins encoded by these genes to continue to function in the plastid, the gene transfer necessitated the development of a post-translational

plastid protein import system, which now is responsible for importing ~95% of chloroplast proteins. These proteins, which are required for chloroplast biogenesis and function, are encoded by genes in the nucleus, translated on cytosolic ribosomes, and post-translationally imported into chloroplasts (Leister 2003).

1.4 Chloroplast Protein Import

The vast majority of the proteins required for plastid biogenesis, function, and interconversion between plastid types is nuclear encoded. The structural and functional variability observed between plastid types is matched by that of their required protein complement. There are approximately 3000 nuclear encoded proteins post-translationally targeted to the chloroplast in *Arabidopsis thaliana* (Leister 2003). While more than one import pathway exists, the majority of nuclear-encoded chloroplast proteins are imported by a proteinaceous, oligomeric apparatus known as the **t**ranslocon at the **o**uter and **i**nner envelope membranes of **c**hloroplasts (T**o**c and T**i**c, respectively- Figure 2) (Schnell et al. 1997; Smith et al. 2006; Inaba & Schnell 2008). Proteins imported via this pathway are translated with a cleavable, amino-terminal targeting sequence, known as a transit peptide, which is necessary and sufficient for import (Bruce 2001). These so-called pre-proteins destined for the chloroplast are recognized at the outer chloroplast envelope membrane by a receptor, and through a series of energy dependent and independent steps, the unfolded protein is translocated across the double envelope membrane (Walker et al. 1996; Ruprecht et al. 2010). Once the pre-protein reaches the stroma, the transit peptide is cleaved by a zinc-binding metallopeptidase, called the stromal processing peptidase (SPP), to yield

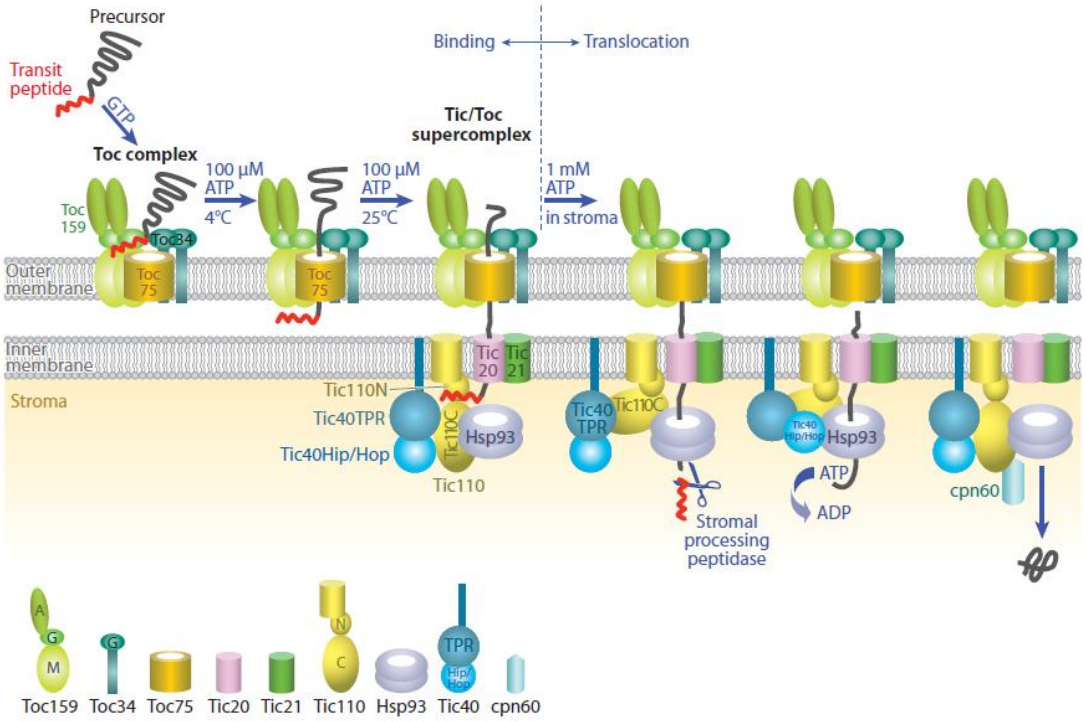


Figure 2. General topology of the Toc-Tic general import apparatus and the proposed sequence of events involved in the import of pre-proteins into the chloroplast (Li & Chiu 2010). This diagram does not include accessory components or intermembrane space components such as Tic22.

either a mature, functional protein or one that is further targeted to other sub-compartments of the chloroplast (Figure 2; Li & Chiu 2010) (Inaba & Schnell 2008).

An increasing number of non-canonical chloroplast proteins that lack a cleavable transit peptide and/or use alternate import pathways are being discovered (Armbruster et al. 2009). Chloroplast envelope Quinone Oxidoreductase Homologue (ceQORH) is targeted to the stromal side of the inner envelope membrane of chloroplasts but lacks a cleavable, N-terminal transit peptide (Miras et al. 2002). It has instead a ~40 amino acid internal sequence that targets this protein to its final destination by an energy dependent system other that the Toc-Tic system (Miras et al. 2007). Tic32 is another chloroplast protein targeted to the inner envelope membrane via an alternative pathway. This protein is a putative regulator of the Toc-Tic import system and while it's targeting information in is located at the N-terminus, this sequence is not removed upon import (Nada and Soll 2004).

A large number of outer envelope membrane proteins, approximately 24, are targeted to chloroplasts via a Toc-Tic independent pathway (Inaba and Schnell 2008). However, the proposed central protein channel, Toc75 is targeted to the outer envelope membrane (OEM) of chloroplasts by a bipartite transit peptide (Tranel et al. 1995; Tranel & Keegstra 1996; Inoue et al. 2001). The N-terminus of the Toc75 transit sequence directs the protein to the chloroplast and is cleaved, like other canonical transit peptides, by SPP (Tranel et al. 1995; Tranel & Keegstra 1996). The C-terminal portion of the transit sequence possesses a polyglycine segment, though not sufficient, is necessary for halting Toc75 at the OEM (Tranel & Keegstra 1996; Inoue et al. 2005; Baldwin & Inoue 2006). Many other OEM proteins, like Toc34 – a

C-terminally tail anchored (TA) protein component of the Toc complex, possess their targeting sequences in and/or adjacent to their transmembrane domains. In a study by Dhanoa et al. (2010), the Toc34 homologues from *Arabidopsis* were shown to target to the OEM of chloroplasts by signaling sequences localized to their GTPase domains and the region adjacent to their transmembrane domains (TMD). In addition, the targeting signal of the OEM protein OEP9 was found in the TMD as well as the adjacent region. It is clear that the central β-barrel channel of the Toc complex, Toc75, is important for the targeting of other OEM proteins; such as Toc159, which requires both Toc75 and Toc34, and OEP7/14, which utilizes Toc75 and a cytosolic chaperone (Wallas et al. 2003; Tu et al. 2004). While evidence of chaperone-mediated pathways is limited, the cytosolic ankyrin-repeat protein, AKR2A, was shown to interact with, and target OEP7/14 to the chloroplast surface and may also be involved in the targeting of the Toc34 homologues (Bae et al. 2008; Dhanoa et al. 2010). There have been recent studies to identify other cofactors, such as Hsp17.8, that affect the AKR2A-mediated targeting of pre-proteins to the chloroplast (Kim et al. 2011). Previously, AKR2A was thought to function in anti-oxidation metabolism and disease resistance (Yan et al. 2002), but the work by Bae et al (2008) suggests its involvement in the chaperone-mediated pathways in chloroplast protein import.

The targeting of proteins to the intermembrane space of chloroplasts has not been studied as extensively as the targeting of proteins elsewhere in the chloroplast. However, two proteins, Tic22, an intermembrane space component of the Tic complex, and a monogalactosyldiacyl glycerol synthase (MGDG synthase), MGD1, have been examined in depth (Kouranov et al. 1999; Vojta et al. 2007a). Both of

these proteins possess an N-terminal transit sequence, and use the Toc complex in some capacity to gain entry to the chloroplast, but differ in their final stages of import and processing to mature proteins (Vojta et al. 2007a). Cytosolic pre-Tic22 is not translocated to the stroma for processing, but bypasses the Tic complex and moves directly to the intermembrane space; whereas the MGD1 pre-protein enters the stroma, at least partially, and is processed by SPP (Kouranov et al. 1998; 1999).

The translocation of inner envelope membrane (IEM) proteins also differs from that of the majority of stromal proteins. While most IEM proteins possess N-terminal transit sequences, secondary signaling sequences determine how the proteins make their way to the inner membrane (Inaba & Schnell 2008). Two distinct pathways have been described in some detail: the stop-transfer mechanism and the conservative sorting (post-import) mechanism (Li & Chiu 2010). The stop-transfer mechanism involves translocation of the pre-protein by the Toc complex and partial insertion through the Tic channel. Here, the recognition of a stop signal, often found in the transmembrane domain (TMD), halts the translocation and the protein moves laterally through the channel into the inner envelope membrane (Viana et al., 2010; Froehlich & Keegstra 2011). The conservative sorting (post-import) mechanism, utilized by Tic110 and Tic40, involves complete translocation of the pre-protein to the stroma where a processed intermediate is formed (Lubeck et al. 1997; Li & Schnell 2006; Tripp et al. 2007). The precursor of atTic40 possesses a bipartite transit sequence comprised of a canonical N-terminal transit peptide required for stromal targeting. Once processed by SPP, the targeting information in the TMD and adjacent N-terminal region, shown to contain a serine/proline (S/P)-rich domain,

plays a critical role in the insertion of atTic40 into the IEM (Tripp et al. 2007). The translocation of Tic110 to the IEM also involves a stromal intermediate, suggesting a similar targeting pathway (Lubeck et al. 1997). In addition, it was found that targeting information C-terminal to the transit peptide was required for localization of Tic110 in the IEM, but that the process was mediated by an undefined stromal, proteinaceous re-export machinery likely involved in other IEM protein targeting (Lubeck et al. 1997; Li & Schnell 2006; Vojta et al. 2007b; Kwoka et al. 2008).

While a number of chloroplast protein import pathways exist, the majority of nuclear encoded chloroplast proteins are translocated by the Toc-Tic import pathway. Despite the extensive research to date on this system, many unanswered questions remain concerning the recognition of pre-proteins, the molecular events of translocation, and the stoichiometry and assembly of the oligomeric protein complexes.

1.5 General Import Apparatus

The two oligomeric protein complexes, Toc and Tic, are located on the outer and inner chloroplast envelope membrane, respectively. The stoichiometry of these complexes at any given time has not yet been elucidated, but there is a model of the general topology of the Toc-Tic components presented in Figure 2. The nomenclature pertaining to proteinaceous components of the Toc-Tic import apparatus as outlined by Schnell et al. (1997) names proteins based on their location; Translocon at the **o**uter and **i**nner envelope membrane of chloroplasts (T**o**c and T**i**c, respectively) with a numerical suffix pertaining to its mass in kilo Daltons. All

components referred to in this work will be derived from the model plant *Arabidopsis thaliana*, unless otherwise stated, and specific members of multi-gene families in *Arabidopsis* carry the prefix "at" to denote their species of origin.

1.5.1 Toc Complex

The Toc complex is comprised of at least five distinct membrane-bound components, three of which represent the core of the complex chiefly responsible for the recognition, binding, and outer envelope membrane translocation of pre-proteins: Toc75, Toc34, and Toc159 (Smith et al. 2006; Inaba & Schnell 2008).

1.5.1.1 Toc75

Toc75 is the most abundant protein in the outer chloroplast envelope membrane and is important for the targeting of other OEM proteins. It is accepted that Toc75 forms at least a portion of the protein translocation channel, since it is capable of forming β-barrel anion channels of 14-26 Å with 16 TMD based on structural analysis; can form cation-selective pores when reconstituted in liposomes, and interacts directly with pre-proteins (Perry & Keegstra 1994; Hinnah et al, 2002). Toc75 is likely to have arisen from the chloroplast endosymbiotic ancestor (Agne et al. 2009) as homology between outer membrane channels of cyanobacteria and other gram negative bacteria suggest a function in protein trafficking (Reumann et al. 2005). Toc75 was identified as a Toc complex component, as determined by cross-linking studies, and interacts directly with another core complex protein, Toc159 (Hinnah et al. 2002; Kouranov & Schnell, 1997). In *Arabidopsis*, Toc75 is a member

of a small multi-gene family consisting of atToc75-I, atToc75-III, and atToc75IV (Jackson-Constan & Keegstra 2001; Eckart et al. 2002). The atToc75-I gene is a pseudogene that is no longer expressed due to a number of mutations and insertions (Baldwin et al, 2005). However, both the atToc75-III and atToc75-IV gene products are expressed *in vivo*, and both encode for integral, β-barrel proteins (Baldwin et al, 2005). The atToc75-III gene product is expressed at higher levels in rapidly dividing tissues, and knockout mutants of this gene are embryo-lethal (Baldwin et al, 2005). In contrast, the atToc75-IV gene is expressed uniformly in most tissues but much lower compared to atToc75-III, and codes for a smaller gene-product that does not require processing or significant energy consumption for insertion, which suggests it plays a role in dark or energy depleted conditions (Baldwin et al, 2005). Like atToc75-III, its paralog, OEP80, is necessary for embryo viability (Hsu & Inoue 2009). The function of OEP80, although undetermined, appears to be distinct from that of Toc75, as OEP80 does not tightly associate with Toc complexes and single knockout mutants of both OEP80 and Toc75 are embryo lethal (Hsu et al. 2012). While it is evident that considerable work is required to determine the membrane orientation and function of OEP80, its impact on chloroplast biogenesis is substantial (Hsu et al. 2012).

1.5.1.2 Toc34

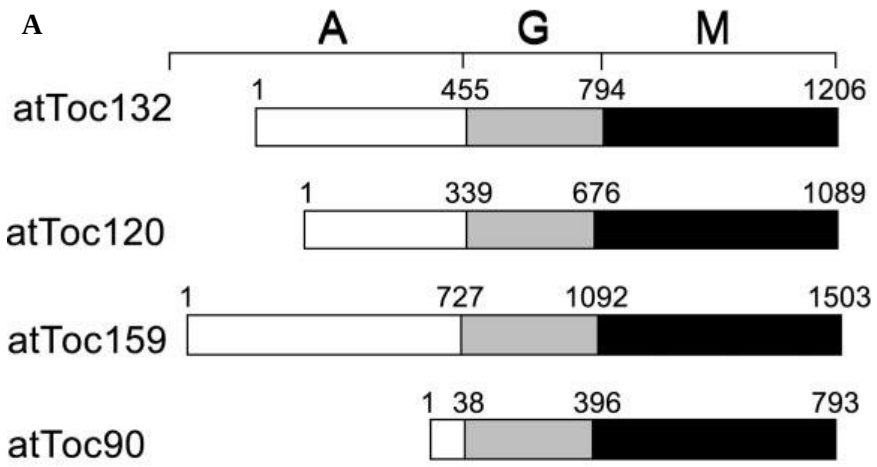
The Toc159 and Toc34 components of the core Toc complex are homologous GTPases that act as the receptors involved in the initial recognition, binding, and insertion of pre-proteins through the outer membrane. In *Arabidopsis*, the Toc34

GTPase is represented by a multi-gene family, atToc33 and atToc34 (Kubis et al. 2003; Constan et al 2004). The Toc34 homologues are tail anchored (TA) proteins which possess a single C-terminal α-helix inserted in the outer chloroplast envelope membrane that protrudes into the intermembrane space, and contain a GTPase (G)-domain exposed to the cytosol (Tsai et al. 1999; Kessler et al. 1994). This GTPase is capable of homodimerization, potentially heterodimerization with Toc159, and binding pre-proteins via the transit peptide (Kouranov et al. 1997; Sun et al. 2002; Weibel et al. 2003; Reddick et al. 2007; Aronsson et al. 2010; Oreb et al. 2011). Previously, the homodimerization of atToc33 was proposed to activate GTPase activity based, in part, on the observation that when dimerized, an arginine residue on the dimer interface was in close proximity to the nucleotide binding site of the opposing atToc33 monomer, similar to that of other known GTPase activating proteins (GAP) (Sun et al. 2002). Testing this hypothesis became difficult because high-affinity dimerization occurred in a nucleotide depleted, or GDP bound state. Jelic et al. (2003) proposed that pre-protein binding could increase GTP hydrolysis of Toc34, acting as a GAP. They conclude from their results that GTP-bound Toc34 was activated upon pre-protein binding, hydrolysis occurred, and the pre-protein was released. The binding of pre-proteins by Toc34 was shown to increase the rate of GTP hydrolysis up to 30-fold (Jelic et al. 2002). However, Oreb et al. (2011) recently demonstrated that pre-proteins actually act as a GDP dissociation inhibitor (GDI) displacement factor (GDF). When atToc33 is in a monomeric state, GTP hydrolysis is free to occur at a constant rate in the presence and absence of pre-proteins, as no nucleotide exchange factors (GEF) or GAPs are required. When atToc33 is in a

dimeric state, the hydrolyzed GDP in the nucleotide binding site of each monomer is closed in by the other and nucleotide exchange can not occur. Using fluorescently labeled GDP analogue, *mant*GDP, the GDP dissociation rates of monomeric and dimeric atToc33 were compared. The results show increasing dissociation rates as the proportion of monomeric atToc33 to dimeric atToc33 increases (Oreb et al. 2011). This result was further supported by using the dimerization deficient mutant atToc33R130A, as an increase in protein concentration, accompanied by an increase in dimerization in wild-type atToc33, did not reduce the rate of mantGDP dissociation. The hypothesis that pre-protein binding by the atToc33 dimer disrupts the dimer and releases GDP was confirmed when monomeric and dimeric *mant*GDP bound atToc33 were incubated with pre-protein and the dissociation rate of the dimeric atToc33 was increased whereas that of the monomeric atToc33 was not (Oreb et al. 2011). The role of this regulation of nucleotide exchange in the mechanism of chloroplast protein import is still unclear (Oreb et al. 2011).

1.5.1.3 Toc159

The Toc159 receptor, and members of the Toc159 family, are tripartite in structure; they consist of a C-terminal membrane (M)-domain that anchors them to the chloroplast outer envelope membrane, a central GTPase (G)-domain homologous to that of Toc34, that is capable of binding pre-proteins, and a large, N-terminal, acidic (A)-domain that extends into the cytosol (Figure 3; Bauer et al. 2000; Ivanova et al. 2004; Smith et al. 2004). While no precise function of the A-domain has been elucidated, it has recently been determined that it is a highly phosphorylated,



B

atToc132

	A	G	M
atToc120	36	93	87
atToc159	11	49	49
atToc90	-	41	32

Figure 3. Domain structure of the four tripartite Toc159 GTPase homologues in *Arabidopsis*. The divisions between the A-, G-, and M-domains of Toc159 homologues based on amino acid sequence (A). Part (B) illustrates the amino acid sequence similarity to atToc132 for each homologue across all three domains (Ivanova et al. 2004).

intrinsically unstructured protein (IUP) domain and there is evidence to suggest that it plays a role in conferring substrate specificity to the Toc159 family of receptors (Richardson et al. 2009; Agne et al. 2010; Inoue et al. 2010). The Toc159 and Toc34 pre-protein receptors in the core of the Toc complex are also represented by multi-gene families in *Arabidopsis* (Bauer et al. 2000; Jackson-Constan & Keegstra 2001). The gene products of Toc34 and Toc159 interact preferentially with particular homologues from the other family to form structurally and functionally distinct complexes that recognize and import various chloroplast pre-proteins with differential specificity (see section 1.6) (Kubis et al. 2004; Ivanova et al. 2004; Smith et al. 2004; Lee et al. 2009; Bischof et al. 2011).

In another recent study, the Toc159 receptor was found to co-sediment with filamentous actin (Jouhet & Gray 2009). The experiments were conducted with unpurified recombinant full-length Toc159 from *Pisum sativum* (psToc159) expressed in *E. coli*. While the full-length protein interacted with actin, the recombinant GTPase-domain did not. It was not directly tested, but the authors suggested that it may be the Acidic-domain of atToc159 that was responsible for the actin interaction (Jouhet & Gray 2009).

1.5.1.3.1 Intrinsically Unstructured Proteins

Intrinsically unstructured proteins (IUP) are proteins that when in an aqueous environment possess little, if any, stable secondary structure (Uversky 2002). They are capable of interconverting between conformations of a disordered state in some

regions of the protein while maintaining some amount of residual structure throughout others. Intrinsically unstructured proteins are estimated to make up for 36 – 63% of the proteome in eukaryotes, such as plants (Gunasekaran et al. 2003). It is hypothesized that the lack of secondary structure of IUPs is essential for their function, and in some instances allows a protein to interact with a very large number of substrates with high specificity and low affinity (Dyson et al. 2002; Dyson et al. 2005; Fink 2005). For example, IUP domains have been found in a number of human transcription factors and proteins involved in cell division (Fuxreiter et al. 2008; Galea et al. 2008). It was also shown that the presence of these IUP domains is important for their function.

Interactions between IUPs and their substrates can be accompanied by a shift from a relatively disordered state to a more defined conformation based upon the structure of the substrate (Dyson et al. 2002; Fuxreiter et al. 2008). A single IUP may be capable of adopting many different conformations based on the substrate and the position of residues important in the interaction (Dyson et al. 2002). This ability lends itself to the vast number of structurally diverse ligands some receptors are capable of recognizing. It has become evident that the intrinsically unstructured nature of some chloroplast protein import components (i.e. Toc159 receptors) may play a crucial role in the recognition and import of pre-proteins (Richardson et al. 2009).

1.5.1.4 Toc Complex Accessory Proteins

Two other membrane bound proteins of the Toc complex, Toc64 and Toc12 have been discovered and investigated more recently. Toc64 was found to crosslink to other Toc complex components. It has also been shown that Toc64 co-fractionates and with other Toc proteins (Sohrt & Soll, 2000; Schleiff et al. 2003b). It was postulated that Toc64 acts as a docking site for pre-proteins targeted to the chloroplast mediated by Hsp90 through a tetratricopeptide repeat (TPR) domain exposed to the cytosol, but that the pre-proteins were released in an energy dependent manner (Qbadou et al. 2006). However, a separate study by Aronsson et al (2007) showed that Toc64 homologues in *Arabidopsis*, atToc64-III, atToc64-V, and atToc64-I, were not essential for chloroplast pre-protein import utilizing triple-knockout mutants. Therefore the precise role of Toc64 remains unclear. Toc12, a J-domain containing protein, was found to co-purify with outer envelope vesicles from pea. Its function was proposed to be involved in the activation of IMS Hsp70 (Becker et al. 2004). However, the psToc12 used in those experiments corresponds to a portion of one DNAJ-J8 protein in pea that is targeted to the stroma (Chiu et al. 2010), and the corresponding *Arabidopsis* homologues all code for atDNAJ-J8 proteins that localize to the stroma. Furthermore, J8 mutants show no defect in chloroplast protein import (Chiu et al. 2010). Therefore the precise function of Toc12 in chloroplast protein import has yet been determined.

1.5.2 Tic Complex

Similar to the Toc complex, the Tic complex is comprised of five core subunits directly involved in the translocation of pre-proteins: Tic22, Tic20, Tic21, Tic110, and Tic40. While the focus of this thesis was not on the Tic complex, each of the components of this complex will be briefly described in the sections that follow.

1.5.2.1 Tic22, Tic20, and Tic21

Tic22 is a small intermembrane space component of the Tic complex that has been shown to interact directly with pre-proteins during import and may play a role in the interaction between Tic and Toc translocons (Kouranov et al. 1998). It is postulated that Tic20 and Tic21 are involved in the formation of the translocon channel in the IEM as they are both small, hydrophobic proteins predicted to form four α-helical TMD and both are essential for translocation across the IEM and for chloroplast biogenesis (Chen et al. 2002; Teng et al. 2006; Kikuchi et al. 2009; Kasmati et al. 2011; Kovacs-Bogdan et al. 2011). Tic20 has also been shown to interact with the transit peptides of pre-proteins (Kouranov & Schnell 1997). While it has been suggested that Tic20 and Tic21 may be important at different stages of development, Tic complexes containing both proteins have been observed (Chen et al. 2002; Kikuchi et al. 2009). Tic20 and Tic21 appear to be structurally similar to the mitochondrial Tim components Tim23 and Tim17, which have been shown to form pores in the same Tim complex, and may be involved in a similar pore forming mechanism (Kovacs-Bogdan et al. 2011).

1.5.2.2 Tic110

Tic110 is the largest and most abundant of the IEM proteins and is essential for chloroplast biogenesis and function as knockout mutants are embryo lethal (Kovacheva et al. 2005). When reconstituted into proteoliposomes, Tic110 has been reported to form a β-barrel and therefore has been hypothesized to form part of the protein channel through the inner envelope membrane (Heins et al. 2002). Alternatively, it has been shown that the portion of the protein reported to form a membrane-embedded β-barrel actually extends into the stroma, and therefore is not capable of forming a large part of the channel *in vivo* (Jackson et al. 1998; Inaba et al. 2003). Indeed, it has been shown that Tic110 possesses two N-terminal, α-helical TMD that may still be involved in the formation of the inner envelope membrane protein channel (Lubeck et al. 1996; 1997; Jackson et al. 1998). The remaining 97.5 kDa C-terminus of the protein is a large globular domain that extends into the stroma (Jackson et al. 1998; Inaba et al. 2003). This globular stromal domain has been found to interact with pre-proteins, as well as recruit Hsp93, a stromal pre-protein chaperone (Neilsen et al. 1997; Inaba et al. 2003). It is thought that Hsp93 acts as a motor to drive pre-proteins through the final stages of import and plays a role in the IEM-targeting re-export machinery. Finally, the ATPase activity of Hsp93 is responsible for the stromal ATP demands of translocation (Kovacheva et al. 2005; Vojta et al. 2007b).

1.5.2.3 Tic40

Tic40 is a co-chaperone whose N-terminus is inserted into the IEM, with the bulk of the protein extending into the stroma. Tic40 possesses a central tetratricopeptide (TPR) repeat followed by C-terminal Hsp70/90-organizing and Hsp70-interacting (Hop-Hip) domains (Bedard et al. 2007). The TPR domain of Tic40 facilitates the release of pre-proteins from Tic110 by interacting directly with the large stromal domain. While there is no evidence of a direct interaction between Hsp93 and pre-proteins, it seems likely that Tic40 mediates some transfer of pre-proteins from Tic110 to Hsp93. The Hop-Hip domain of Tic40 has also been shown to interact with Hsp93 to stimulate its ATPase activity, providing the energy required to complete translocation (Chou et al. 2006; Bedard et al. 2007).

1.5.2.4 Tic Complex Redox Components

There also exists a group of redox proteins associated with the Tic complex: Tic32, Tic62, and Tic55 (Caliebe et al. 1997; Hormann et al. 2004; Stengel et al. 2008). It is possible that these proteins are involved in the redox-mediated regulation of ferredoxin-III and ferredoxin-NAD(P)$^{+}$ (Hirohashi et al. 2001; Yan et al. 2006; Inaba & Schnell 2008). Tic32 is necessary for chloroplast biogenesis, as knockout mutants in *Arabidopsis* were embryo lethal (Hormann et al. 2004). This may be related to the calmodulin-binding activity of Tic32, which suggests it may play a role in calcium regulation of chloroplast protein import (Hormann et al. 2004).

1.5.3 Transit Peptides

Transit peptides are short extensions of 20-100 amino acids that are typically void of acidic residues and generally enriched in hydrophobic and hydroxylated residues and are found on the N-terminus of the majority of nuclear-encoded proteins destined for the chloroplast (refer to section 1.4) (Bruce 2001). While no functional consensus sequences or motifs have been defined among these highly variable chloroplast targeting sequences, there has been some success with computer predictions for transit peptides (Emanuelsson et al. 1999; Bannai et al. 2002; Small et al. 2004). Evidence exists for the post translational phosphorylation of some transit peptides that stimulates the assembly of a guidance complex composed of chaperones and precursors, which ultimately interacts with the Toc complex. However, the precise nature and importance of this guidance complex remains undefined (May et al. 2000; Smith et al. 2006).

Transit peptides belong to the IUP family of proteins, similar to the A-domain of Toc159 homologues (Wienk et al. 1999). This likely has significant implications for the recognition of multiple, highly variable transit peptides by a single Toc complex as IUPs are capable of gaining conformation when in the presence of an interaction partner (Dyson et al. 2002).

1.6 Multiple Chloroplast Import Pathways

As previously discussed in sections 1.5.1.2 and 1.5.1.3, each of the two GTPase receptors in *Arabidopsis* are represented by multi-gene families. The Toc159 family has four members (atToc159, atToc132, atToc120, and atToc90), and the

Toc34 family has two (atToc33 and atToc34). The sequence identity observed between the G-domains of Toc34 homologues is 61% (Jarvis et al. 1998). The M- and G-domains of atToc132 and atToc120 share approximately 90% sequence identity and share only 50% sequence identity with atToc159 (Ivanova et al. 2004). While functional differences have been observed between these domains despite the high degree of structural similarity observed between Toc34 homologues and the M- and G-domains of relevant Toc159 homologues, it is the 36% sequence identity observed between the A-domains of atToc132 and atToc120, and the 11% between atToc159 and atToc132 that strongly suggests the possibility that the A-domains provide a functional distinction (Ivanova et al. 2004). It has been suggested that the distinct members of the Toc GTPases assemble into structurally distinct Toc complexes that could form for the purpose of recognizing and importing specific subsets of proteins (e.g. photosynthetic vs. non-photosynthetic – Figure 4) (Bauer et al. 2000; Ivanova et al. 2004; Kubis et al. 2004; Inoue et al. 2010; Li & Chiu 2010).

The expression of atToc159 is approximately 10-fold higher than atToc132/120 in photosynthetic tissues, whereas atToc132/120 are reported to be expressed at uniform low levels in all tissue types (Kubis et al. 2004). While this suggests functional differences between homologues, the case for structurally distinct Toc complexes exhibiting differences in functional specificity was strengthened by the use of knockout mutants. Specifically, the knockout mutant of atToc159, **p**lastid **p**rotein **i**mport mutant **2** (*ppi2*), is albino and seedling lethal (Bauer et al. 2000). Upon examination of the plastids, they were found to be void of chlorophyll, thylakoids, and typical photosynthetic proteins, such as small and large subunits of

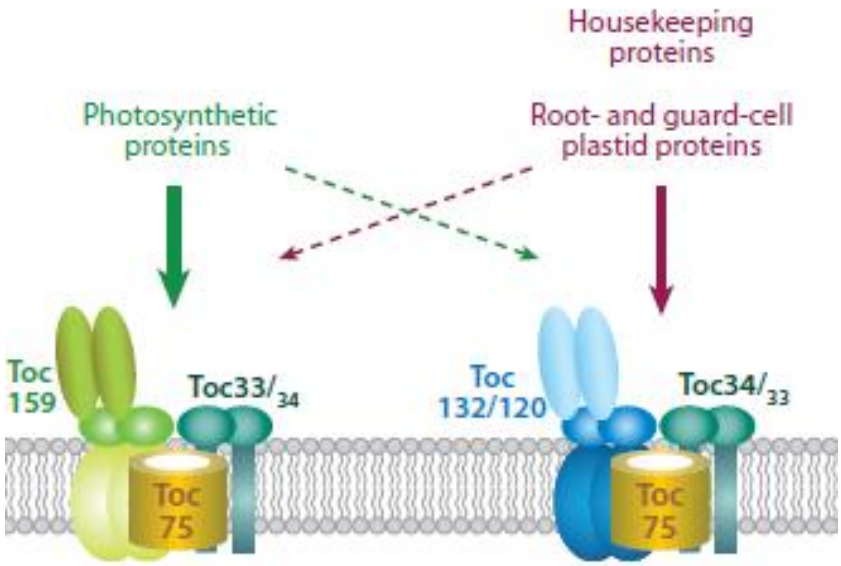


Figure 4. Structurally and functionally distinct Toc complexes. AtToc159 preferentially associates with atToc33 to form Toc complexes to import photosynthetic proteins and atToc132 preferentially associates with atToc34 to form Toc complexes to import non-photosynthetic pre-proteins (Li & Chiu 2010).

Rubisco and chlorophyll *a/b* binding protein (Bauer et al. 2000). While the complementation of atToc159 rescues the phenotype in *ppi2* plants, over expression of atToc132 does not, demonstrating that atToc159 and atToc132/120 containing Toc complexes are functionally distinct (Ivanova et al. 2004). Specific interactions between atToc159 and photosynthetic pre-proteins have also been observed using *in vitro* binding assays (Smith et al. 2004). In addition, the import of a transiently expressed photosynthetic transit peptide-GFP fusion protein (small subunit of Rubisco (prSSU) TP-GFP) into *ppi2* chloroplasts was less efficient than that of a transiently expressed non-photosynthetic transit peptide-GFP fusion (E1α subunit of pyruvate dehydrogenase (prE1α) – GFP) suggesting that atToc159 is required for the import of photosynthetic pre-proteins (Smith et al. 2004). Together, these data suggest that atToc159 is necessary for chloroplast biogenesis and important for the accumulation of photosynthetic proteins. AtToc90 exhibits expression patterns similar to that of atToc159, but in significantly lower amounts, and possesses the same conserved G- and M-domains. Infanger et al. (2011) showed that an over expression of atToc90 partially rescued the *ppi2* phenotype demonstrating slightly elevated chlorophyll, partially recovered leaf anatomy, and a rudimentary thylakoid membrane. This suggests that atToc90 is capable of importing at least some of the pre-proteins that are normally handled by atToc159. In particular, atToc90$_{His}$/ *ppi2* chloroplasts were able to, at a reduced capacity, import pSSU (Infanger et al. 2011).

The knockout mutants of the highly similar atToc132 and atToc120 demonstrate their redundancy. In single knockout plants of either homologue, growth defects were not detected and chlorophyll levels were not noticeably reduced, but

double knockouts were not viable suggesting that some redundancy exists but at least one homologue is required (Ivanova et al. 2004; Kubis et al. 2004). The role of atToc132/120 import of non-photosynthetic pre-proteins was further tested using *in vitro* binding assays in which atToc132 bound a non-photosynthetic pre-protein (prE1α) but not the photosynthetic pre-protein (prSSU) (Ivanova et al. 2004). Moreover, Lee et al. (2009) demonstrated that prE1α-GFP accumulated in chloroplasts of *ppi2* protoplasts, but prSSU fusion import was decreased.

Similar to the Toc159 homologues, atToc33 and atToc34 exhibit differing expression profiles. AtToc33 is abundant in photosynthetic, rapidly dividing tissues and atToc34 is expressed at low, uniform levels throughout most tissues (Jarvis et al. 1998; Gutensohn et al. 2000; Kubis et al. 2003). The Toc34 homologues have also been assessed utilizing knockout mutants that demonstrate functional differences between them. The atToc33 knockout mutant, *ppi1*, exhibits a similar, yet less severe phenotype to the *ppi2* mutant (Jarvis et al. 1998; Bauer et al. 2000; Kubis et al. 2003). Silencing of the atToc33 gene causes a pale phenotype, associated with the significant down-regulation of key photosynthetic genes and reduced protein import specific of photosynthetic pre-proteins and not non-photosynthetic proteins (Kubis et al. 2003). In contrast, the atToc34 knockout mutant, *ppi3*, displays photosynthetic tissue growth similar to wild-type but a reduced root system despite normal root plastid morphology (Constan et al. 2004). Despite the differences observed between these two mutants, the over expression of atToc34 is able to rescue *ppi1* plants (Jarvis et al. 1998). The successful complementation of *ppi1* with atToc34 and the survival of both mutants

suggest some functional redundancy among Toc34 homologues (Jarvis et al. 1998; Constan et al. 2004).

The observed defects in knockout mutants and differential expression patterns exhibited by homologues of both Toc159 and Toc34 suggest that assembly of structurally distinct Toc complexes required for the import of specific subgroups of pre-proteins (e.g. photosynthetic and non-photosynthetic). To verify this, Ivanova et al. (2004) conducted co-immunoprecipitation assays to determine which Toc34 homologue associated with each Toc159 homologue. These studies showed that while Toc75 was present in all complexes containing a Toc159 homologue, complexes that were immune-precipitated with α-atToc159 did not contain atToc132 or atToc120 and vice versa. However, complexes precipitated with α-atToc132 contained both atToc132 and atToc120, providing additional evidence for the functional redundancy of these two homologues (Ivanova et al. 2004). This study also demonstrated that Toc159 homologues preferentially associate with particular Toc34 homologues. Despite a small amount of cross-reactivity, atToc33 precipitated primarily with atToc159-containing complexes and the majority of precipitated atToc34 was found in atToc132/120-containing complexes (Ivanova et al. 2004). The cross-reactivity exhibited by the Toc34 antibody agrees with the previously observed sequence similarity between atToc33 and atToc34 (Jarvis et al. 1998).

The various stages of plant development and adaptation to various environmental stresses require a dynamic, ever changing gene expression profile and the plastid translocon that imports the large number of highly variable proteins needs to be able to effectively sort and prioritize substrates. The structurally and

functionally distinct Toc complexes formed in *Arabidopsis* that are capable of distinguishing between photosynthetic and non-photosynthetic proteins, coupled with the known and as-of-yet undiscovered cytosolic factors and regulatory mechanisms are well suited to dealing with the dynamic chloroplast proteome that requires post-translational import. However, questions still remain concerning the nature of the primary pre-protein receptor and the interactions between the Toc GTPases in an active Toc complex.

1.7 Pre-protein Recognition by atToc159 – Role of the A-domain

Functionally distinct Toc complexes capable of recognizing a large number of different pre-protein substrates and differentiating between sub-categories of those substrates would require a primary receptor that is capable of those same things. The preference of atToc33 and atToc34 to form separate structurally and functionally distinct Toc complexes suggests that there are structural differences between these two homologues sufficient for the formation of these complexes despite the 61% sequence identity they share (Jarvis et al. 1998; Kubis et al. 2003). It is not likely however, that these relatively small components of distinct Toc complexes possess pre-protein binding sites with the plasticity and substrate specificity to recognize the large number of distinct and highly variable transit peptide sequences that target pre-proteins to chloroplasts.

The sequence identity between the Toc159 homologues atToc159 and atToc132/120 (only atToc132 will be considered from this point on due to functional redundancy) is considerably lower than the identity between atToc33 and atToc34;

~50% identity between M- and G-domains and 11% between A-domains (Ivanova et al. 2004). The lower sequence identity observed between G-domains would suggest that the pre-protein binding sites of atToc159 and atToc132 could differ more in their substrate specificity than those between Toc34 homologues. The most striking difference between atToc159 and atToc132 is the differences observed in the functionally uncharacterized A-domains. Consider the size of the A-domain relative to the total size of the protein; the A-domains of atToc159 and atToc132 account for approximately 48% and 38% of the total protein, respectively (Ivanova et al. 2004). The total length of atToc159 and atToc132 is different by 297 amino acids, and their respective A-domains differ by 272 amino acids, sharing only 11% sequence identity (Ivanova et al. 2004). Discounting the reduced sequence identity observed between the M- and G-domains of these Toc159 homologues, the extreme differences observed between the A-domains suggests that the Toc159 component of distinct Toc complexes could be acting as the primary pre-protein receptor, or regulating the primary receptor's interactions with pre-proteins.

As previously discussed, from analysis of *ppi2* chloroplasts, it can be inferred that atToc159 is important for the import of photosynthetic pre-proteins (Bauer et al. 2000). The ability of atToc159 to bind the transit peptides of the precursor to small subunit of Rubisco (prSSU) and precursor to ferredoxin (prFD), model photosynthetic proteins, in a dose-dependent manner by the G-domain was confirmed by Smith et al (2004) using *in vitro* binding assays. AtToc159 was also shown not to bind non-photosynthetic precursors or mature proteins lacking transit peptides (Smith et al. 2004).

In a recent study, Inoue et al. (2010) demonstrated that the A-domain of the Toc159 components of distinct complexes is responsible for their ability to distinguish between photosynthetic and non-photosynthetic pre-proteins. The pre-protein receptors, atToc159 and atToc132 were shown to bind model photosynthetic (ferredoxin (Fd)) and non-photosynthetic (E1α) transit peptide fusions, respectively, in a dose-dependent manner that could not be competed by a transit peptide of the opposite sub-class when full length receptors possessing all three domains were used. However, when mutant atToc159 and atToc132 lacking their respective A-domains were used (atToc159GM and atToc132GM, respectively), the binding of respective target proteins was competed by both photosynthetic and non-photosynthetic transit peptides in the case of both receptors. To follow these observations, the import of the prFD fusion competed by the prE1α fusion was observed in chloroplasts isolated from wild-type, full length atToc159 rescued *ppi2* (atToc159/*ppi2*), and mutant atToc159GM rescued *ppi2* (atToc159GM/*ppi2*). The results from the wild-type and atToc159/*ppi2* experiments were indistinguishable and demonstrated that import of the prFD fusion could only be competed by itself and not the prE1α fusion. The import of the prFD fusion into atToc159GM/*ppi2* chloroplasts was competed by prFD in a manner similar to that observed in wild-type, but E1α transit peptides also reduced import to ~40% of wild-type (Inoue et al. 2010). This study provides substantial evidence to suggest that atToc159 is a selective pre-protein import receptor whose specificity is mediated by the A-domain or that the A-domain of atToc159 confers specificity for another receptor in the Toc complex.

As mentioned in section 1.5.1.3.1, the A-domains of the Toc159 homologues belong to the IUP family of proteins that are capable of interacting with a large number of substrates with high specificity (Gunasekaran et al. 2003; Richardson et al. 2009). While it is clear that the A-domain does not interact directly with pre-protein transit peptides *in vitro*, it is likely that the mechanism of pre-protein binding specificity conferred by the A-domain is complex. It is possible that a multi-protein interaction involving the A-domain mediates the recognition of precursors at the Toc complex (Smith et al. 2004). *In vivo*, it has also been observed that the A-domain exists as a hyperphosphorylated protein domain (Agne et al. 2010). In the same study, the proteolytic susceptibility of this protein domain, generally assumed to be a sensitivity not a functional characteristic, was observed *in vivo*. Even more interesting was the localization of atToc159A expressed on its own in transgenic plants to both the soluble and chloroplast membrane fraction of isolated tissue (Agne et al. 2010). The large size, significant potential for regulation by phosphorylation, and unstructured conformation of the A-domain allows for considerably more solvent-exposed residues capable of interacting with binding partners. The role of the A-domain in pre-protein binding specificity is further substantiated by the IUP nature of positively charged transit peptides, as the A-domain, when unphosphorylated, is predicted to carry a charge of -136 at physiological pH (Putnuam, The Scripps Research Institute).

1.8. Protein-Protein Interactions

Biological responses at the molecular level such as signal transduction, enzyme activation, and import/export pathway stimulation are often the result of protein-ligand or protein-protein interactions. Indeed, the Toc and Tic complexes involve extensive protein-protein interactions among the constituents of the complexes. A number of possible interactions between compatible residues of proteins including hydrogen bonding, Van der Waals interactions, ionic interactions, and covalent binding can occur alone or in combination to produce unique changes to the structure and/or character of the proteins involved. Due to the large number of possible outcomes of protein-protein interactions, there are many ways to detect and measure them.

1.8.1 Circular Dichroism Spectropolarimetry

Circular dichroism (CD) is a spectroscopic technique that examines and characterizes common secondary and tertiary structures of proteins using spectropolarimetry (Kelly et al. 2000). In CD spectroscopy, a beam of right and left-handed circularly polarized light is passed through the sample. Differential absorption occurs at chiral centers, such as those found in the peptide backbone, which absorb the right and left-handed light differently (Greenfield, 2007; Kelly et al. 2000). The transmitted beam that exits the sample is no longer circularly polarized, but elliptically polarized.

There are two major wavelength regions used for protein structure elucidation during CD spectroscopy: far UV region (180 nm – 240 nm) and the near UV region (260

nm– 320 nm) (Figure 5). The far UV region is used to examine secondary structure motifs. The examination of data from this region will be the main focus of the experiments in the current study (Kelly et al. 2000). CD spectra are converted to molar ellipticity, which normalizes measurements based on the concentration and number of chiral centers found in a particular sample. Molar ellipticity measures the differential absorption of the circularly polarized light, which results in elliptically polarized light, per amino acid in solution (Kelly et al. 2000). This ensures that the absorption can be normalized based on concentration of amino acids so that peptides with very different molecular weights but similar structure will yield similar CD spectra.

1.9 Overall Objectives

There is significant evidence to suggest that the A-domain of Toc159 homologues, namely atToc159 and atToc132, are involved in conferring the specificity of structurally and functionally distinct Toc complexes (Smith et al. 2004; Inoue et al. 2010). The IUP nature of the A-domain of Toc159 strongly suggests a role in pre-protein recognition as IUP domains are capable of interacting with a large number of substrates with high specificity (Dyson et al. 2005). The large proportion of the protein that the A-domain accounts for and the extreme variation in this domain between homologues would allow atToc159 and atToc132 to contribute to the functional specificity observed in their respective Toc complexes. The objective of the current study was to further characterize the nature of the role played by the A-

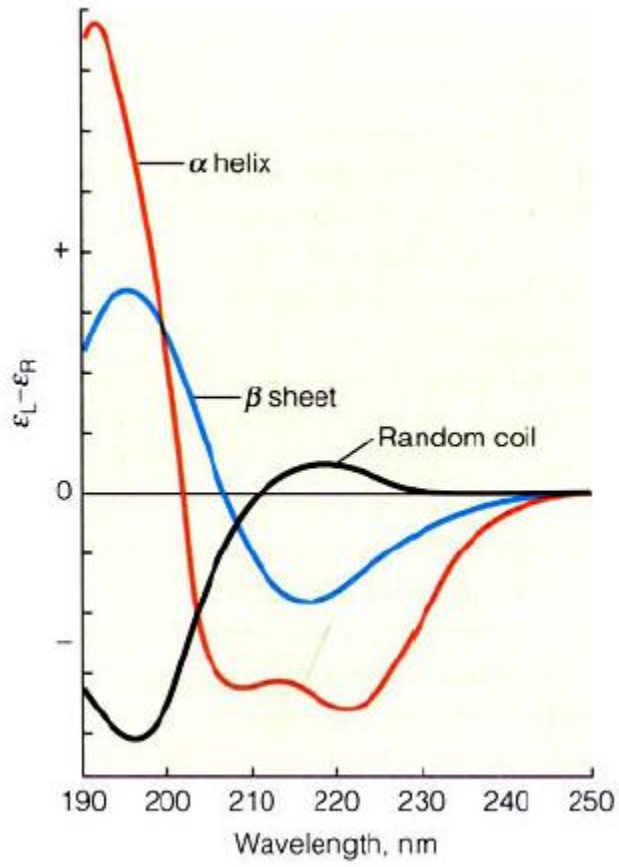


Figure 5. CD spectra of characteristic secondary structures measured in the Far UV range (Mathews 1996).

domain of atToc159 in conferring specificity to photosynthetic pre-protein importing complexes, and the overall function of Toc159 and its homologues.

1.9.1 Hypothesis

The large size, high sequence variability among homologues, and intrinsically unstructured nature of the A-domain, coupled with the previous observation that the A-domain of Toc159 does not directly interact with transit peptides is in agreement with the A-domain's ability to confer pre-protein binding specificity by distinct Toc complexes. Based on this, it was hypothesized that the Acidic-domain interacts with the GTPase-domain of atToc159 to confer the functional specificity observed between structurally distinct Toc complexes by regulating the Toc159 GTP hydrolysis involved in pre-protein import.

1.10 Specific Aims

In order to further characterize the role of the A-domain in chloroplast protein import, this study focuses on the protein-protein interactions observed between the A-domain and the GTPase domain of atToc159. A number of different approaches have been used in an attempt to functionally characterize the A-domain of atToc159 and atToc132. Limited proteolysis of wild-type chloroplasts with thrombin, which cleaves atToc159 close to the division between the A- and G-domains, was utilized to determine whether atToc159A is associated with other components the chloroplast outer envelope membrane. To expand on the thrombin experiments, *in vitro*-translated atToc159A was targeted to wild-type and A-domain deficient,

atToc159GM$_{His}$ / *ppi2* chloroplasts. Solid-phase binding assays were used to determine which Toc complex components the Toc159 A-domains interacted with. To examine the observed interactions further, circular dichroism spectropolarimetry and GTP hydrolysis assays incorporating both the A- and G-domains of atToc159 were conducted. In response to a study by Jouhet & Gray (2009) that observed atToc159 interacting with actin, the possibility that filamentous actin interacts with the A-domain of atToc159 was also tested.

2. Materials and Methods

2.1 Cloning of atToc132A Construct

A construct coding for an untagged protein product corresponding to the Acidic domain of atToc132 was required for generating *in vitro* transcription-translation products for solid-phase pull-down experiments.

Primers were designed to amplify the A-domain of atToc132 (1-1365 nt from the cDNA, Ivanova et al. 2004) from a construct containing full length atToc132 (pET21a:atToc132FL, Ivanova et al. 2004) and introduce a 5' **NdeI** restriction site atToc132ANoHis.NdeI.S - (5'-ATAGTTATA**CATATG**GGAGATGGGACCGAG-3'), a 3' **XhoI** restriction site, and a *STOP* codon atToc132ANoHis.XhoI.AS - (5'-TTCTA**CTCGAG**T*CTA*AAGACCTGCTGGACG-3'). The atToc132A fragment was amplified by PCR using HotStar Hifidelity Polymerase (Qiagen). The 50 µl PCR reaction was set up according to the manufacturer's specifications: 2.5 U of polymerase, 25 ng of template (pET21a:atToc132FL), primers at 1 µM each. The 35-cycle program was also based on the manufacturer's suggestions: 1 minute annealing phase at 58^0C, 1 minute 23 second (1 min/kb) extension phase at 72^oC. Two 50 µl PCR reactions were pooled and purified using a PCR Clean-up Kit (Promega). The purified atToc132A PCR product and 4 µg of pET21a were digested with 40 U of NdeI (NEB) at 37^oC for 2.5 and 3.5 hours, respectively. The digested insert was purified using PCR Clean-up Kit (Promega) and the digested pET21a was gel purified (Promega). Both the insert and vector, eluted in 50 µl of water, were digested with 40 U of XhoI (NEB) at 37^oC for 2.5 and 3 hours, respectively; the insert and vector were subjected to a final PCR Clean-up (Promega). The insert was ligated with the

digested pET21a using MightyMix Ligation Mix (TAKARA BIO) at a molar ration of 5:1, atToc132A (101 ng) : pET21a (100 ng), ratio at 16^{o}C for 30 minutes. The ligation product was used to transform DH5α *E. coli* using electroporation (see section 2.1.1). Colonies were screened for the presence of the atToc132A insert by colony PCR (Refer to Colony PCR Screen section – Appendix I). Positive colonies were cultured, plasmids were isolated and the sequence of the insert was confirmed (TCAG Sequencing, Sick Kids Hospital, Toronto) using sequencing primers corresponding to the T7 Promoter and T7 Terminator sequences on the pET21A vector. Glycerol stocks were made from positive colonies carrying inserts with the correct sequence.

2.1.1 Transformation of *E. coli*

For the construct containing the cDNA of atToc132A (pET21a:atToc132A$_{NoHis}$), 20 ng of plasmid (~1.5 µl of ligation mixture) was mixed with 100 µl of electrocompetent *E. coli* (DH5α or BL21 RIPL) and incubated on ice for 60 seconds. The mixture was transferred to an electroporation cuvette (BioRad) and electroporated using the "Ec1" bacterial setting (Bio-Rad MicroPulser). Cells were allowed to recover at 37^{o}C, shaking at 240 RPM for 1 hour. Cells were then plated on LB-agar (DH5α: 50 µg/ml ampicillin; BL21 RIPL: 50 µg/ml chloramphenicol, 75 µg/ml streptomycin). Cultures were made in LB (plus appropriate antibiotics) from individual positive colonies identified by colony PCR and stored at -80^{0}C in 15% glycerol.

2.2 Expression of Recombinant Proteins in BL21 RIPL

The following protocol was used for the expression of atToc159A_{His} (pET21d:atToc159A_{His}, Richardson et al (2009), atToc159G_{His} (pET21d:atToc159G_{His}, Smith et al. 2004), atToc132A_{His} (pET21a:atToc132$_{His}$, Richardson et al 2009), atToc132G_{His} (pET21a:atToc132G_{His}, Farquharson, W. unpublished; Chen M.Sc. 2011), atToc33G_{His}, and atToc34G_{His} (pET21d:atToc33G_{His} and pET8C:atToc34G_{His} respectively, Weibel et al. 2003) in E. coli BL21 RIPL (Stratagene). Overnight cultures of 4 ml, inoculated with the a loop of culture from glycerol stock, were grown in the presence of the required antibiotics (50 µg/ml ampicillin, 50 µg/ml chloramphenicol, 75 µg/ml streptomycin) for 16-18 hours. 1 liter of LB broth containing 50 µg/ml ampicillin and 50 µg/ml chloramphenicol was inoculated with 4 ml of overnight culture. The 1 liter culture was grown at 37^{o}C with shaking at 240 RPM until an OD_{600} of 0.6-0.8 was reached. Expression was induced with the addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Induction of the A-domain cultures was achieved at ~22^{o}C with shaking at 180 RPM for 16-18 hours; the G-domain cultures were induced at 37^{o}C with shaking at 240 RPM for 3.5 hours. Cells were collected by centrifugation at 8000xg at 4^{o}C for 15 minutes in a JA10.50 rotor (Avanti J-30I, Beckman Coulter). The supernatant was discarded and cell pellets were stored at -20^{o}C.

2.3 Purification of atToc159A$_{His}$, atToc132A$_{His}$ by Immobilized Metal Affinity Chromatography (IMAC)

The IMAC purification of atToc159A$_{His}$ and atToc132A$_{His}$ was adapted from Richardson et al (2009). Frozen cell pellets were resuspended in 40 ml of Ni-159A-B (10 mM Tris-HCl, pH 8.0, 300 mM NaCl, 5 mM imidizole). Cells were lysed by incubating for 45 minutes at 4^{o}C while rotating in 0.8 mg/ml lysozyme (Egg white – Bioshop) and sonicating on ice using a probe-tip sonicator, manually pulsed on and off for 10 minutes. Insoluble components of the cell lysate were removed by centrifugation at 20000xg, at 4^{o}C for 60 minutes in a JA 30.50 rotor (Avanti J-30I, Beckman Coulter). A column (Flex Column – 1.5x15cm, Kimble-Chase) containing 2 ml of Ni-NTA His-Bind resin (Novagen) was washed with 20 ml of sterile water followed by two washes with 10 ml of Ni-159A-B. Soluble lysate containing the his-tagged recombinant protein was applied to the column, allowed to flow through, and collected. The column was washed with 20 ml of Ni-159A-B followed by 12 ml of Ni-159A-W (10 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10 mM imidizole). Proteins were eluted in 1 ml fractions with Ni-159A-E (10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 250 mM Imidizole) and stored in 10% glycerol at -80^{0}C. The presence of protein was confirmed by resolving samples on SDS-PAGE gels (4% stacking, 10% resolving) under constant voltage (110V). Proteins were visualized by staining with Coomassie blue (0.5% (w/v) Coomassie Blue R250, 50% (v/v) methanol, 10% (v/v) acetic acid).

2.3.1 Purification of atToc159A$_{His}$, atToc132A$_{His}$ by Ion Exchange Chromatography

The ion exchange purification of atToc159A$_{His}$ and atToc132A$_{His}$ was adapted from Richardson et al (2009). Multiple IMAC-pure elution fractions were combined and diluted 1:1 in IEx-159A-B (20 mM piperazine pH 4.5, 200 mM NaCl). 2 ml of Q Sepharose Fast Flow ion exchange resin (GE Healthcare) was transferred to a 20 ml glass screw cap tube and washed four times with 15 ml of IEx-159A-B. The 1:1 diluted protein sample was applied to the resin and allowed to incubate at room temperature on a rotator for 10 minutes. The supernatant was removed and the resin was washed twice with 15 ml of IEx-159A-B. 5 ml of IEx-159A-E (20 mM piperazine pH 4.5, 550 mM NaCl) was applied to the resin and allowed to incubate on a rotator for 10 minutes. The supernatant was removed from resin and centrifuged using a bench top centrifuge to remove all resin for buffer exchange.

2.3.2 Buffer Exchange by Centrifugal Filtration

A 15 ml Amicon 10 kDa limit centrifugal filter (Millipore) was equilibrated with 15 ml of CD-159 (10 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 50 mM NaCl, 1 mM DTT) by centrifugation at 5000xg for 20 minutes in a JA 30.50 rotor. All of the ion exchange-purified atToc159A$_{His}$ was applied to the filter and diluted to 15 ml with CD-159. Once applied to the filter, the samples were centrifuged at 5000xg and 4^{o}C for 20 minutes in a JA 30.50 rotor (Avanti J-30I, Beckman Coulter). The retentate volume was recorded and the flow through was discarded. The retentate volume was corrected to 15 ml with CD-159 and centrifuged. This step was repeated twice more

and retentate volumes were recorded for determination of the imidazole concentration of the final sample. Presence of the proteins following the final concentration step was confirmed using SDS-PAGE (4% stacking, 10% resolving) under constant voltage (110V) and stained with Coomassie blue (0.5% (w/v) Coomassie Blue R250, 50% (v/v) methanol, 10% (v/v) acetic acid). Protein concentration was estimated by BioRad Protein Assay.

2.4 Purification of atToc159G$_{His}$, atToc33G$_{His}$, atToc132G$_{His}$, atToc34G$_{His}$ by Immobilized Metal Affinity Chromatography (IMAC)

The IMAC purification of atToc159G$_{His}$, atToc132G$_{His}$, atToc33G$_{His}$, and atToc34G$_{His}$ were adapted from Smith et al (2004); Chen (2011). The frozen cell pellet was resuspended in 40 ml of Ni-159G-B (10 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10 mM Imidizole). Cells were lysed using lysozyme (Egg white, Bioshop) at 0.8 mg/ml, rotating at 4^{o}C for 45 minutes and manually probe sonicated for 10 minutes. Cell lysates were cleared by centrifugation at 20 000xg, 4^{o}C for 45 minutes in a JA 30.50 rotor (Avanti J-30I, Beckman Coulter).

2 ml of Ni-NTA His-Bind resin (Novagen) in a small Flex Column (1.5x15 cm, Kimble-Chase) was washed with 20 ml of sterile water followed by two washes with 10 ml of Ni-159G-B. Sample was applied to the column, the flow-through was collected, and the resin was washed with 20 ml of Ni-159G-B followed by 12 ml of Ni-159G-W (10 mM Tris-HCl, pH 8.0, 300 mM NaCl, 40 mM Imidizole). Proteins were eluted in 1 ml fractions with Ni-159G-E (10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 250 mM Imidizole) and stored in 10% glycerol at -80^{0}C. The presence of

protein was confirmed by resolving samples in SDS-PAGE gels (4% stacking, 12% resolving) under constant voltage (110V) and stained with Coomassie blue (0.5% (w/v) Coomassie Blue R250, 50% (v/v) methanol, 10% (v/v) acetic acid). Protein concentration of elutions was estimated using BioRad Protein Assay.

2.5 Protein Concentration Estimation by Bio-Rad Microassay

All protein concentration estimations were achieved using this method. A 10 mg/ml bovine serum albumin (BSA) stock was used for the standards. Six standards were used with concentrations ranging from 0-10 µg/ml. Dilutions of each protein sample to be quantified were made in 800 µl volumes. 200 µl of Bio-Rad Protein Assay Dye Reagent concentrate was added to each sample and allowed to incubate at room temperature for 30 min. Each sample was transferred to a 1 cm, polycarbonate cuvette and absorbance at 595 nm was measured in triplicate for each sample using the multicell Cary50 UV Spectrophotometer.

2.6 Interaction of atToc159A and atToc159G by CD Spectroscopy

An Aviv 215 Circular Dichroism Spectropolarimeter was used for all CD measurements. Samples of atToc159G_{His} containing imidazole were dialyzed 1:300 against CD-159 Buffer (10 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 50 mM NaCl, 1 mM DTT) for 3.5 h at 4^{o}C. All samples were measured in a 0.01 cm quartz cuvette in the far-UV range (190 nm - 260 nm); 4 scans at 0.5 nm/s at intervals of 0.5 nm. For each concentration and condition, at least two independent samples were prepared to 55 µl in CD-159 and centrifuged briefly prior to measurement. Measurements of buffer

and both proteins, with and without nucleotide (GTP) were taken at multiple concentrations alone and in combination.

Data analysis was completed using the Aviv 215 software and Origin 7.0. The average of 8 scans corresponding to each duplicate sample was taken and the average of the buffer blank was subtracted. All resulting spectra were automatically smoothed and mean residue ellipticity was calculated based on protein concentration and number of amino acids in each protein. Samples of individual proteins were measured and their spectra were added (see Appendix II). The addition of the individual spectra was compared to the spectra corresponding to a mixed protein sample to determine if any changes in secondary structure occurred upon mixing of the two proteins. The protein concentrations in the mixed samples were not equimolar in all cases and required weighting based on number of amino acids and concentration to calculate molar ellipticity. The concentration of atToc159A$_{His}$ was used as a baseline for this as it remained constant in all samples.

$$\#\ \text{Amino Acids} = \left[\frac{[\text{atToc159G}_{His}]}{[\text{atToc159A}_{His}]}\right] \bullet (\#\ \text{A.A. atToc159GHis}) + (\#\ \text{A.A. atToc159AHis})$$

2.7 Solid Phase Pulldown Assays

The solid phase pull-down assay protocol was adapted from Smith et al. (2002). Briefly, 200 µl of Ni-NTA His-Bind resin (Novagen) was resuspended in 400 µl water and centrifuged at 10000xg, for 30 sec at room temperature. After the supernatant was removed, this washing protocol was repeated twice more. The resin was then washed using the same protocol with HMK (50 mM HEPES-KOH, pH 7.5,

5 mM $MgCl_2$, 40 mM KOAc, 1 mM THP) in triplicate. After the resin was HMK-equilibrated, it was diluted with 5 volumes of HMK. The slurry was distributed into 12, 40 µl aliquots to which HMK and His-tagged "bait" protein (atToc159G_{His}, atToc33G_{His}, atToc132$G_{His,}$ or atToc34G_{His}) were added to a final volume of 500 µl in the following amounts, in duplicate: 0, 100, 200, 300, 400, 500 pmol. The bait was immobilized by incubation on a rotator at room temperature for 30 min. The immobilized bait was re-equilibrated in HMKITG (50 mM HEPES-KOH, pH 7.5, 5 mM $MgCl_2$, 40 mM KOAc, 10 mM Imidazole, 0.1% Triton X-100, 0.1 mM GTP, 1 mM THP) by pelleting the resin, removing the supernatant, and washing, as previously described above, with 250 µl HMKITG once. Once re-equilibrated, 2 µl or 3 µl of radiolabeled "prey" protein, atToc159A_{NoHis} or atToc132A_{NoHis}, respectively, were added to each tube with 98 µl of HMKITG and incubated on a rotator at room temperature for 30 min. The tubes were centrifuged at 10000xg, for 30 sec at 4^{o}C and the supernatant was discarded. The samples were then washed as outlined above three times with 400 µl of ice cold HMKITG at 4^{o}C. Once washed, the proteins were eluted in 20 µl SBI (0.17 M Tris base, 1% (w/v) bromophenol blue, 2.5% (w/v) SDS, 80 mM DTT, 3.75% glycerol, 775 mM imidazole). See section 2.15.1 for sample analysis.

2.8 GTPase Activity Assay

GTP hydrolysis of recombinant atToc159G_{His} and atToc159A_{His} was measured using a method adapted from Smith et al. (2002). Proteins, alone and in combination, were diluted in GTPase Buffer (50 mM HEPES-KOH, pH 7.5, 5 mM

$MgCl_2$, 40 mM KOAc, 50 mM NaCl, 1 mM DTT) to a concentration of 5 mM (each) in 30 µl. The protein was nucleotide depleted by incubating at 25°C for 30 min. Hydrolysis was initiated by the addition of GTP (975 nM unlabeled GTP (Sigma), 25 nM [α^{32}P]-GTP (Perkin Elmer) and reactions were incubated at 25°C. The reaction was stopped by denaturation at 0, 5, 15, 30, and 60 min time points by removing 5 µl of the reaction and adding to 10 µl of Stop Solution (0.5% SDS, 10 mM EDTA, 4 mM GTP, 4 mM GDP) and heating at 65°C for 2 min. Following denaturation, 2 µl of each sample was spotted onto 20 cm PEI-Cellulose F TLC plates (EMD) and resolved using 1 M LiCl solvent. The plates were dried and placed in an autoradiography cassette and exposed to a phosphor-screen for 60 min. Following exposure, the screen was scanned at 100 µm resolution (Personal Molecular Imager FX, BioRad). The detected spots of radioactivity, corresponding to radiolabeled GTP and GDP were quantified using Quantity One 1-D Analysis software. The amount of hydrolysis at a given time point was determined as a percentage of atToc159G_{His} activity calculated as a ratio of GDP to GTP as compared to the zero time point of each reaction.

2.9 Growth of *Arabidopsis thaliana*

2.9.1 Plant growth for propagation

Plants were propagated on soil for seed harvest. Both wild-type (Columbia ecotype) and atToc159GM_{His} rescued *ppi2*/WT plants (*ppi2*:atToc159GM_{His}, kindly provided by Dr. Danny Schnell, U. Massachusetts) were placed on wet soil, 5 seeds

per pot, and grown at 21^{o}C under a 16 h :8 h, light:dark cycle in an environment-controlled growth chamber (Bigfoot, Econair).

Mutant plants were treated with BASTA (Glufosinate ammonium, Sigma) to select for transformants. Following a 5 day growth period after germination, mutants were sprayed with a 300 µM BASTA solution every 3 days for 2 weeks. Plants were grown until flowering ceased and then allowed to dry out for seed harvest.

2.9.2 Growth of *Arabidopsis thaliana* Plants for Chloroplast Isolation

The protocol for the growth of *Arabidopsis thaliana* for isolation of intact chloroplasts was adapted from Fitzpatrick & Keegstra (2001) and Smith et al (2003). Seeds were surface sterilized by washing in 95% ethanol for 5 minutes followed by washing for 30 min in 30% bleach containing 0.02% (v/v) Triton-X 100. In a sterile flow hood, seeds were rinsed 5 times with sterile water. Approximately 30 mg of seeds were sown onto 150 mm x 15 mm plates (0.5x Murashige and Skoog media, 1% sucrose, 0.8% Phytagar). AtToc159GM$_{His}$ rescued *ppi2*/WT seeds were sown onto plates supplemented with 50 µg/ml kanamycin (Bioshop) to select for *ppi2* transformants.

All plates were sealed with Parafilm and plants were grown for 16-21 days at 21^{o}C under 16 h:8 h, light:dark cycle in a controlled growth chamber (Bigfoot, Econair).

2.10 Isolation of Intact Chloroplasts from Plate-grown *Arabidopsis*

Chloroplast isolation protocol was adapted from Schulz et at. (2004) and Brock et al. (1993), as described by Richardson (2009). Two Percoll step gradients

were prepared in round-bottom centrifuge tubes (Nalgene) by layering 8 ml of 35% Percoll (10 mM HEPES-KOH, pH 7.5, 2 mM EDTA, 1 mM $MgCl_2$, 1 mM $MnCl_2$, 330 mM sorbitol, 50 mM ascorbic acid, 0.05% plant protease inhibitor cocktail (P9599, Sigma), 35% Percoll (v/v) (Sigma)) on 7 ml of 85% Percoll (components the same as 35% Percoll solution, Percoll increased to 85% (v/v)). Gradients and other solutions were chilled on ice prior to isolation. Green tissue was harvested from 3-5, 16-21 day-old plates using a razor blade. Tissue was immersed in 150 ml of fresh 1x Grinding Buffer (10 mM HEPES-KOH, pH7.5, 2 mM EDTA, 1 mM $MgCl_2$, 1 mM $MnCl_2$, 330 mM sorbitol, 100 mM ascorbic acid, 0.25% (w/v) BSA, 0.05% plant protease inhibitor cocktail (P9599, Sigma)) and homogenized using a PowerGen Homogenizer (Fisher Scientific) for 11 seconds. Homogenate was filtered through two layers of Miracloth (Calbiochem) into a pre-chilled 500 ml centrifuge bottle (Beckman Coulter). The sample was centrifuged at 1000xg, for 8 minutes at 4^oC in a JLA 10.500 rotor (Avanti J-30I, Beckman Coulter). The supernatant was discarded and the pellet was gently resuspended in 8 ml of 1x Grinding Buffer shaking on ice. Once suspension was uniform, half of the suspension was layered onto each of the percoll gradients and soft-spun at 7700xg, 4^oC for 15 minutes in a JS13.1 rotor (Avanti J-30I, Beckman Coulter). The top layer, containing broken chloroplasts, was removed by aspiration. The second layer, containing intact chloroplasts, was removed using a Pasteur pipette and diluted into 20 ml of ice cold HS buffer (50 mM HEPES-KOH, pH 7.5, 330 mM sorbitol) in a round-bottom centrifuge tube and brought to ~50 ml with ice cold HS buffer. Dilute, intact chloroplasts were centrifuged at 1000xg, 4^0C for 6 min in a JS13.1 rotor. The supernatant was discarded and the

chloroplast pellet was gently resuspended in 200-300 µl of ice cold HS buffer until the suspension was uniform.

Chloroplasts were quantified by concentration of chlorophyll as described by Arnon (1949). The resuspended chloroplasts were diluted 100-fold in 80% acetone and the absorbance was measured at 652 nm (Cary50 UV-Vis Spectrophotometer). The sample was quantified using the following formula:

$$[Cp]\ mg/ml = \frac{A_{652} \times \text{Dilution Factor}}{36} \qquad \text{Dilution Factor} = 100$$

For targeting assays and proteolysis assays, intact chloroplasts were diluted to 1 mg/ml in ice cold HS buffer.

2.11 Limited Proteolysis of Wild-Type (WT) Chloroplasts by Thrombin

Protocol adapted from Inoue and Schnell (unpublished, personal communication). Approximately 300 µl of 1 mg/ml wild-type chloroplasts were prepared. In duplicate, 4 reactions were prepared in 1.7 ml microfuge tubes containing increasing amounts of Thrombin (0, 0.1, 0.2, and 1 units (U), Sigma) – which proteolytically cleaves atToc159 between P756 and N757 - in HS buffer (50 mM HEPES-KOH, pH 7.5, 330 mM sorbitol) to a volume of 15 µl; one set of reaction tubes were pre-incubated at 26^{o}C, the other on ice, for 5 min. After pre-incubation, 25 µl of 1 mg/ml chloroplasts (equivalent to 25 µg total chlorophyll) was added to each reaction tube and incubated for 30 min . All samples were centrifuged at 4000 RPM, for 3 min at 4^{o}C. The supernatant was transferred to a fresh tube and

held on ice, and each chloroplast pellet was resuspended in 20 µl of 2xSB (0.35 M Tris base, 2% (w/v) bromophenol blue, 5% (w/v) SDS, 0.16 M DTT, 7.5% glycerol).

To TCA precipitate the total protein in the supernatant fraction, 4.4 µl of 100% trichloroacetic acid (TCA) was added to each of the supernatant fractions and incubated, on ice, for 4 h. Samples were centrifuged at 14000 RPM, 4^{o}C for 60 minutes. The supernatant was removed; the pellet was washed with 500 µl of 0.5% TCA and centrifuged again at 14000 RPM, 4^{0}C for 10 minutes. The supernatant was removed and the pellet was resuspended in 20 µl of 2xSB by room temperature water bath sonication.

2.11.1 Western Blotting – Thrombin Treated WT Cp

The chloroplast pellets and TCA precipitated pellets resuspended in 2xSB were resolved by SDS-PAGE (4% stacking, 8% resolving) at constant voltage (110V). The gels were soaked briefly in Transfer Buffer (12.5 mM Tris-HCl, pH 7.5, 96 mM glycine, 0.05% (w/v) SDS, 10% (v/v) methanol) prior to transfer to 0.2 µm-pore nitrocellulose membrane (Whatman) at 15V for 120 min using a Trans-Blot SD Semi-Dry transfer apparatus (Bio-Rad). The membrane was rinsed thoroughly in water and lightly stained with amido black (45% (v/v) methanol, 10% (v/v) acetic acid, 0.1% (w/v) amido black) to ensure complete transfer; the positions of the standards in the protein ladder (SDS6H2, Sigma) were marked with pen. The membrane was washed thoroughly with water and blocked overnight in 5% (w/v) powdered milk in TBS-T (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% (v/v) Tween-20) shaking overnight at 4^{0}C. The membrane was washed twice, shaking in

TBS-T for 5 min at room temperature. The membrane was incubated with primary antibody, rabbit raised α-atToc159A (a generous gift from Dr. Danny Schnell, U. Massachusetts; Dr. Felix Kessler, U. Neuchatel, Switzerland), diluted 1:2000 in TBS-T, 1% (w/v) BSA for 2 h rotating at room temperature. The membrane was washed in TBS-T three times for 5 min each, while shaking at room temperature. The membrane was then incubated with the secondary antibody, peroxidase-conjugated goat anti-rabbit IgG (Rockland), diluted 1:5000 in TBS-T, 1% (w/v) BSA for 2 h rotating at room temperature. To ensure the removal of all weakly bound and unbound secondary antibody, the membrane was washed in TBS-T 5 times for 5 min each at room temperature. Once removed from solution, 10 ml of Luminata Crescendo Chemiluminescence solution (Millipore) was poured evenly over the membrane and allowed to incubate for 5 min. The solution was poured off and the membrane was imaged using a VersaDoc4000 (Bio-Rad) on ChemiSuper mode, with a AF Nikkor 50 mm lens (Nikon), exposed for 60-500 seconds with an f-stop of 1.4. The chemiluminescence signal was analyzed using Quantity One 1-D Analysis software (Bio-Rad).

2.12 *In Vitro* Translation of Radiolabeled Acidic Proteins

Constructs corresponding to the un-tagged Acidic domains of atToc159 (pET21d:atToc159A_{NoHis}) and atToc132 (pET21a:atToc132A_{NoHis}) or precursor to small subunit of Rubisco (prSSU) (pET21a:AtpS1B) were isolated from *E. coli* DH5α to be used as templates for the *in vitro* transcription and translation of the corresponding proteins. The radiolabeled atToc159A_{NoHis} and prSSU products were

generated using the T_NT Coupled Reticulocyte Lysate System (Promega) and the atToc132A_{NoHis} product was generated using the T_NT Coupled Wheat Germ Extract System (Promega). In each case, [^{35}S]-Methionine (EasyTag EXPRESS ^{35}S Protein Labeling Mix, Perkin Elmer) was added to the reaction with methionine-deficient amino acid mix, 1 µg of template DNA, and the other kit components as per the manufacturer's recommendations. Reactions were incubated at 30^{o}C for 90 min and placed on ice prior to use. It is of note that these particular radiolabeled products degrade when stored at -80^{o}C and were therefore prepared fresh for each experiment.

2.13 *In Vitro* Chloroplast Targeting

I*n vitro* chloroplast targeting assays were conducted as described by Smith et al (2002b). Freshly isolated, intact chloroplasts were prepared for each assay as described in section 2.10. Each in vitro chloroplast targeting reaction contained 50 mM HEPES-KOH, pH 7.5, 1 mM DTT, 1 mM GTP, 5 mM ATP, 10 mM methionine, 330 mM sorbitol, 5 mM MgOAc, 40 mM KOAc, and intact chloroplasts equivalent to 50 µg chlorophyll brought to 95 µl with HS buffer (50 mM HEPES-KOH, pH 7.5, 330 mM sorbitol). Reactions were incubated at 26^{0}C for 5 min before 5 µl of radiolabeled atToc159A_{NoHis} or atToc132A_{NoHis} were added. Reactions were incubated at 26^{o}C, in the light, for 30 min. Following incubation, each reaction was stopped by dilution into 400 µl of ice cold HS buffer. The stopped reaction in HS buffer was layered on top of 800 µl of ice cold 35% Percoll and centrifuged at 6000xg, for 6 minutes at room temperature. The top layer of broken chloroplasts and most of the percoll was removed. The remaining sample was transferred to a clean

microfuge tube containing 800 µl of ice cold HS buffer supplemented with 2 mM EDTA, vortexed and incubated on ice for 10 min. The sample was centrifuged at 14000 RPM, 4^{o}C for 30 min. The supernatant was removed and the pellet resuspended in 20 µl of 2xSB (350 mM Tris base, 2% (w/v) bromophenol blue, 5% (w/v) SDS, 0.16M DTT, 7.5% glycerol). See section 2.15.2 for sample analysis.

2.14 Actin Co-sedimentation

Actin was converted to monomeric state by diluting 25 µg of human non-muscle actin (referred to as actin) (Cytoskeleton Inc) to 0.4 mg/ml with Actin Buffer (5 mM Tris-HCl, pH 8.0, 0.2 mM $CaCl_2$, 0.2 mM ATP, 0.5 mM DTT) and incubating on ice for 60 min. Sample was centrifuged at 13 000 xg for 15 min. Supernatant was transferred to a fresh tube and protein concentration was estimated by BioRad Protein Assay. 5x Polymerization Buffer (250 mM KCl, 5 mM ATP, 10 mM $MgCl_2$) was added to monomeric actin to 1x concentration and incubated at room temperature for 60 min to produce filamentous actin (F-actin). A sample of atToc159A_{His} was centrifuged at 100 000 xg, 22^{o}C for 20 min in the TLA 100.3 rotor (Optima TLX Ultracentrifuge - Beckman Coulter). Reactions containing 100 pmol of F-actin were set up with or without 50 pmol of atToc159A_{His} to 150 µl in Actin Rxn Buffer (10 mM Tris-HCl, pH 7.5, 1 mM ATP, 1 mM GTP, 5 mM $MgCl_2$, 0.2 mM $CaCl_2$, 1 mM DTT) and incubated at room temperature for 60 min. All reactions were centrifuged at 100 000 xg, 22^{o}C for 45 min. Supernatant was removed from pellet and TCA precipitated (see section 2.11.1). Pellets containing F-actin and any co-sedimented protein were resuspended in 20 µl 2xSB (350 mM Tris base, 2% (w/v) bromophenol

blue, 5% (w/v) SDS, 0.16M DTT, 7.5% glycerol). See section 2.15 for sample analysis.

2.15 SDS-PAGE and Phosphor-imager Analysis

The recombinant A-domains purified by IMAC and ion exchange and samples from actin co-sedimentation assays were resolved by by SDS-PAGE (4% stacking, 10% resolving) at constant voltage (110V) along side SDS 6H2 protein marker (Sigma). The recombinant G-domains purified by IMAC were resolved by SDS-PAGE (4% stacking, 12% resolving) at constant voltage (110V) along side SDS 7 protein marker (Sigma)

The membrane-associated proteins in chloroplast pellets isolated in the final step of *in vitro* chloroplast targeting assays and total eluted protein from solid-phase pull-down assays were resolved by SDS-PAGE (4% stacking, 12% resolving) at constant voltage (110V) along side a sample of *in vitro* translated radiolabeled protein (10% of the amount used in the assay). All SDS-PAGE gels were stained with Coomassie R-250 (Bioshop). The gels without *in vitro* translated [^{35}S]-labeled protein were imaged using VersaDoc4000 (Bio-Rad) on Coomassie mode, with a AF Nikkor 50 mm lens (Nikon), an F-stop of 5.6, and a 0.1 s exposure time. Gels with *in vitro* translated [^{35}S]-labeled protein were dried onto filter paper under vacuum at 80^{0}C using the gel drier (BioRad). The dried gels were enclosed in autoradiography cassettes and exposed to a phosphor-screen for 3-5 days. After sufficient exposure, the screen was scanned at 100 µm resolution (Personal Molecular Imager FX, BioRad). The detected radioactivity, corresponding to radiolabeled proteins targeted to the

chloroplast or bound by immobilized bait, was quantified using Quantity One 1-D Analysis software (Bio-Rad). Percent-bound protein, to either chloroplast membrane or bait protein was calculated based on the 10% reference sample.

3. Results

Functionally distinct Toc complexes that contain distinct Toc159 homologues and recognize specific sub-groups of pre-proteins posses an unknown mechanism of specificity. The IUP nature, large size, and high variability of the A-domain suggests that it plays a role in conferring that observed specificity (Ivanova et al. 2004; Richardson et al. 2009). In a study by Inoue et al. (2010), the A-domain was shown to confer the specificity observed between structurally and functionally distinct Toc complexes. However, the mechanism by which the A-domain facilitates the Toc complex's distinction between photosynthetic and non-photosynthetic pre-proteins is largely unknown, although its IUP nature suggests a role in directly mediating the interaction between transit peptides and the primary receptor (see section 1).

3.1 The A-domain of atToc159 associates with chloroplasts when cleaved by thrombin.

To test whether the A-domain might interact with other components of the chloroplast, wild-type chloroplasts were treated with increasing amounts of thrombin, which cleaves atToc159 between residues P756 and N757 (Smith et al. 2004), to remove the A-domain from the rest of the atToc159 protein. Re-isolated chloroplasts and associated supernatants were analyzed on Western blots probed with α-atToc159A. The full length atToc159 present at the chloroplast (Figure 6, lane 1) was increasingly degraded as the amount of thrombin increased, but a band corresponding to the A-domain of atToc159 is observed in both the soluble (Figure 6, lanes 2, 4, 6, 8) and chloroplast fraction (Figure 6, lanes 1, 3, 5, 7).

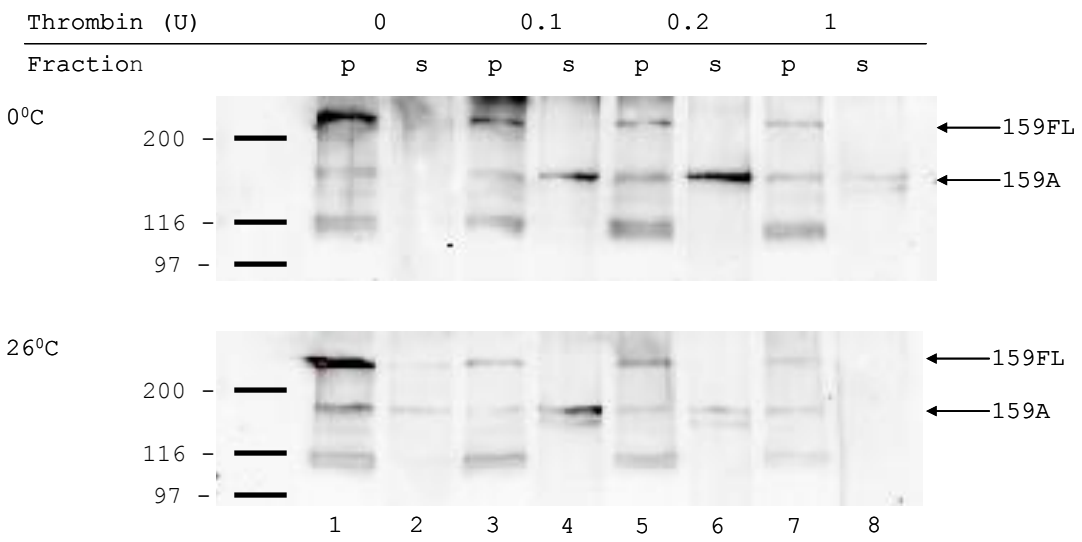


Figure 6. Limited proteolysis of wild-type chloroplasts with thrombin. The experiment was run on ice (0^{o}C – top panel) and at 26^{o}C (26^{o}C – bottom panel) with 0, 0.1, 0.2, and 1 unit (U) of thrombin. Thrombin-treated WT chloroplast samples were analyzed as separate chloroplast (lanes 1, 3, 5, 7) and soluble (lanes 2, 4, 6, 8) fractions by Western blotting and probed with α-atToc159A. The amount of thrombin in each reaction is labeled at the top of the figure in units (U). The bands on the blots corresponding to full length atToc159 (159FL) and the Acidic-domain of atToc159 (159A) are also labeled as identified based on their migration compared to the protein marker (kDa).

The association of the A-domain of atToc159 is most evident in samples that were treated with thrombin at 0^{o}C, as degradation is limited as compared to the treatment at 26^{o}C.

3.2 *In vitro*-translated A-domain of atToc159 does not target to wild-type or A-domain-deficient chloroplasts *in vitro*

Reticulocyte lysate- and wheat germ lysate-derived [^{35}S]-atToc159A were incubated with wild-type and atToc159GM$_{His}$/*ppi2* chloroplasts to determine whether or not the A-domain is able to bind to the chloroplast outer membrane. A faint band corresponding to <1% of reticulocyte lysate-derived [^{35}S]-atToc159A loaded associated with wild-type chloroplasts (Figure 7, B-lanes 2, 3), but there was no evidence of successful targeting of atToc159A to A-domain deficient chloroplasts (Figure 7, C, D).

3.3 *In vitro* binding of the A-domains of atToc159 and atToc132 by GTPase-domain binding partners.

The observation that a portion of the A-domain remains associated with chloroplasts after proteolytic cleavage by thrombin opens the question of what the A-domain interacts with. One of the characteristics of IUPs, such as the A-domain, is that they can interact with a large number of binding partners. The most likely candidate in this case is another component(s) of the Toc complex, including other domains of atToc159 itself. In addition, there are reports in the literature of IUP

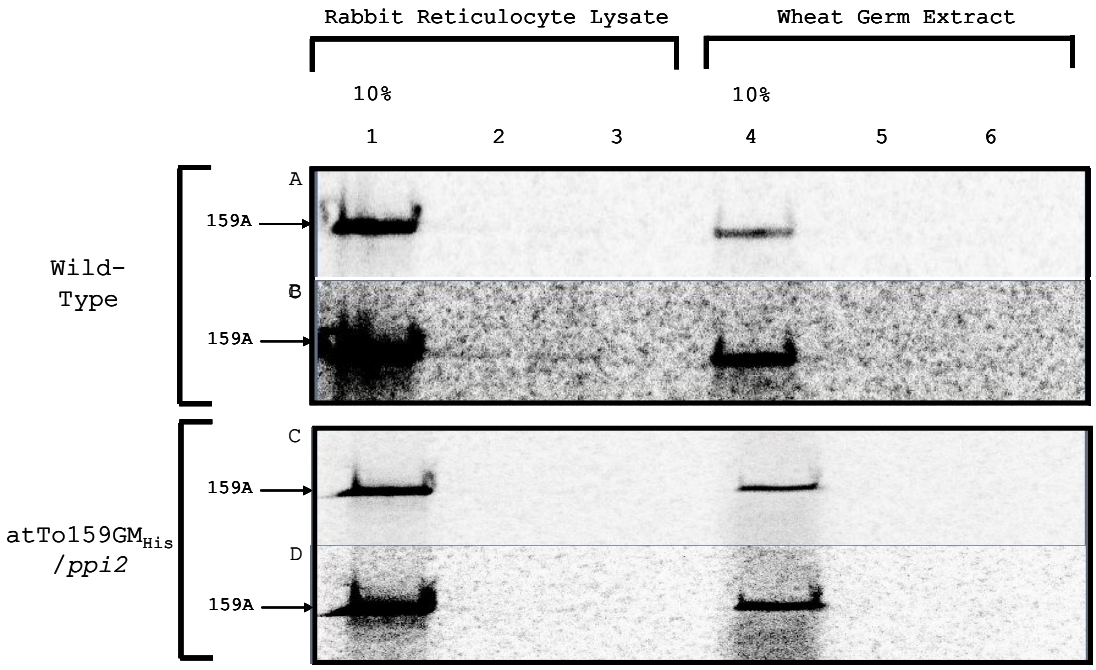


Figure 7. *In vitro* targeting of atToc159 A-domain to chloroplasts. The A-domain of atToc159 associated weakly to wild-type (A, B) and did not associate with A-domain-deficient chloroplasts (C, D) *in vitro*. The Reticulocyte lysate (lanes 1-3) and wheat germ extract (lanes 4-6) derived [^{35}S]-atToc159A was targeted to wild-type (WT, A, B) and atToc159GM$_{His}$ /*ppi2* chloroplasts (C, D). Samples were resolved using SDS-PAGE, and radiolabeled 159A was detected in dried gels using a phosphor-imager. Lanes 1 and 4 correspond to 10% of the translated product used in each assay. The remaining lanes correspond to SDS-PAGE-resolved total chloroplast fractions re-isolated from targeting assays. Panel B is an over-exposed version of panel A and panel D is an over-exposed version of panel C ,to visualize low levels of targeting.

domains interacting with, and modulating the activity of, other enzymatic domains of the same protein (Jonker et al. 2006). To determine whether or not the A-domains of the Toc159 homologues interact with the corresponding GTPase domains of the same protein or other Toc GTPase domains, solid-phase binding assays were used. C-terminally hexa-histidine tagged Toc GTPase domains corresponding to atToc33G_{His}, atToc34G_{His}, atToc132G_{His}, and atToc159G_{His} were used as "bait" proteins and [^{35}S]-labeled atToc159A and atToc132A as "prey" proteins. The A-domain of atToc159 interacted with the GTPase-domain of atToc159 in a dose-dependent manner, but did not interact with the GTPase domains of atToc132, or atToc33 (Figure 8, A). Conversely, the A-domain of atToc132 interacted with the GTPase-domain of atToc132 in a dose-dependent manner, but did not interact with the GTPase domains of atToc159, or atToc34 (Figure 8, C). The highest level of A-domain binding that was observed for atToc159G was ~15% of loaded prey at 250 pmol of atToc159G (bait), whereas maximal binding of the atToc132 A-domain to the atToc132 GTPase domains was ~11% of loaded prey at 500 pmol of atToc132G.

To test whether or not the interaction between the A-domain and G-domain of Toc159 homologues was competitive, we attempted to produce an unlabeled, recombinant A-domain that lacked a poly-His purification tag. However, the development was problematic and is ongoing so the competition portion of these assays was not completed.

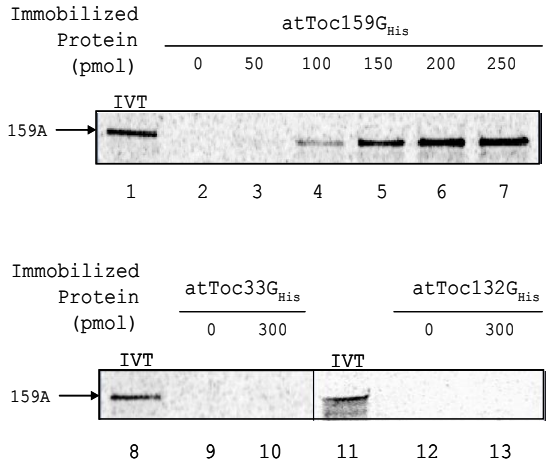
3.4 Influence of the A-domain on GTPase activity of the atToc159 G-domain.

The hydrolysis of GTP by the recombinant GTPase-domain of atToc159 has been measured previously (Smith et al. 2002). In the current study, activity was monitored over 60 min, and confirmed that the recombinant protein was correctly folded and active (Figure 9). The activity was calculated first based on the ratio of GDP to GTP, resolved by TLC and detected by phosphor-imaging, and then an effort to functionally characterize the observed interaction between atToc159 A- and G-domains, the GTPases activity of atToc159G was measured when combined with an equimolar amount of atToc159A. The presence of the A-domain decreased the level of activity when compared to the activity of atToc159G alone by ~50% (Figure 9; C, D).

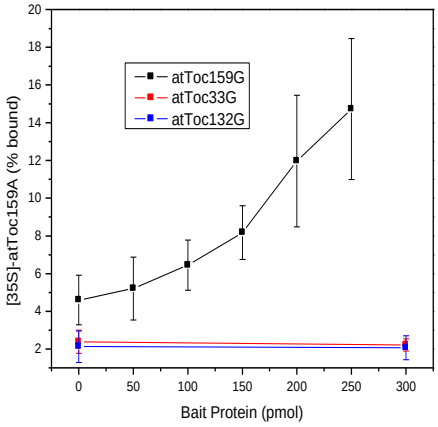
3.5 A- and G-domains of atToc159 measured by circular dichroism spectropolarimetry

The CD spectra of atToc159A_{His} (12.5 μM) (Figure 10, A) and atToc159G_{His} (25 μM) (Figure 10, C) were measured from 190 nm to 260 nm in the presence and absence of 1 mM GTP. Deconvolution software was used to estimate the secondary structure content in each case (Figure 10, B, D). The spectra of atToc159A_{His} with and without GTP (Figure 10, A) are very similar, as is reflected in the negligible changes in secondary structure elements (Figure 10, B). The spectra of atToc159G_{His} (Figure 10, C) with and without GTP exhibited notable differences. The addition of GTP caused an increase in intensity of the maxima at 195 nm and the minima at 208 nm and 222 nm, suggesting an increase in α-helical character associated with the binding of GTP.

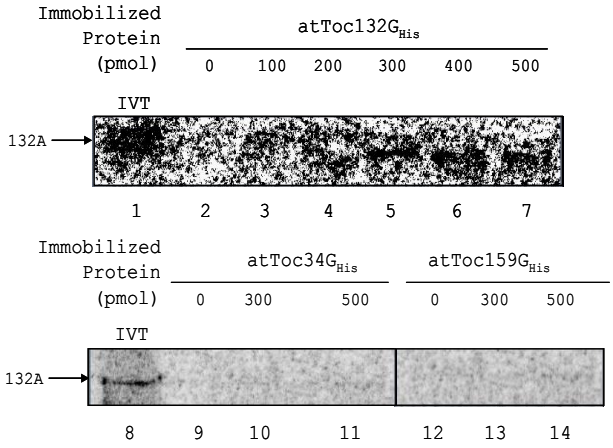
A


Immobilized Protein (pmol)
atToc159G$_{His}$
0 50 100 150 200 250
IVT
159A
1 2 3 4 5 6 7
Immobilized Protein (pmol)
atToc33G$_{His}$
atToc132G$_{His}$
0 300
0 300
IVT
IVT
159A
8 9 10 11 12 13

B


atToc159G
atToc33G
atToc132G
[35S]-atToc159A (% bound)
Bait Protein (pmol)

C


Immobilized Protein (pmol)
atToc132G$_{His}$
0 100 200 300 400 500
IVT
132A
1 2 3 4 5 6 7
Immobilized Protein (pmol)
atToc34G$_{His}$
atToc159G$_{His}$
0 300 500
0 300 500
IVT
132A
8 9 10 11 12 13 14

D

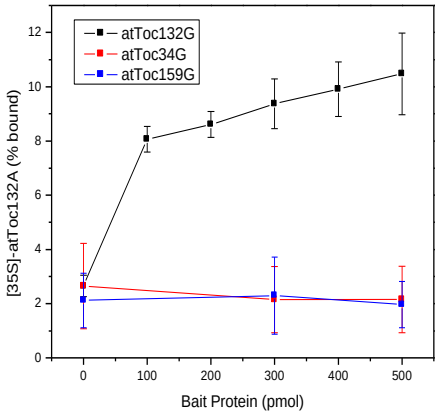

atToc132G
atToc34G
atToc159G
[35S]-atToc132A (% bound)
Bait Protein (pmol)

Figure 8. A-domains of Toc159 homologues specifically interact with their corresponding GTPase domains. [^{35}S]-atToc159A was incubated with increasing amounts of immobilized atToc159G$_{His}$, atToc33G$_{His}$, or atToc132G$_{His}$ (Figure 8, A). [^{35}S]-atToc132A was incubated with increasing amounts of immobilized atToc132G$_{His}$, atToc34G$_{His}$, or atToc159G$_{His}$ (Figure 8, C). The phosphor-image of [^{35}S]-atToc159A (Figure 8, A – lanes 2-7, 9-10, 12-13) and [^{35}S]-atToc132A (Figure 8, C – lanes 2-7, 9-14) bound by immobilized prey together with the 10% reference sample (IVT) (Figure 8, A - lanes 1, 8, 11, C – lanes 1, 8) are shown. Binding is represented as a percentage of total IVT prey loaded in each reaction in the corresponding plot (B and D, respectively). Data was collected from three independent experiments.

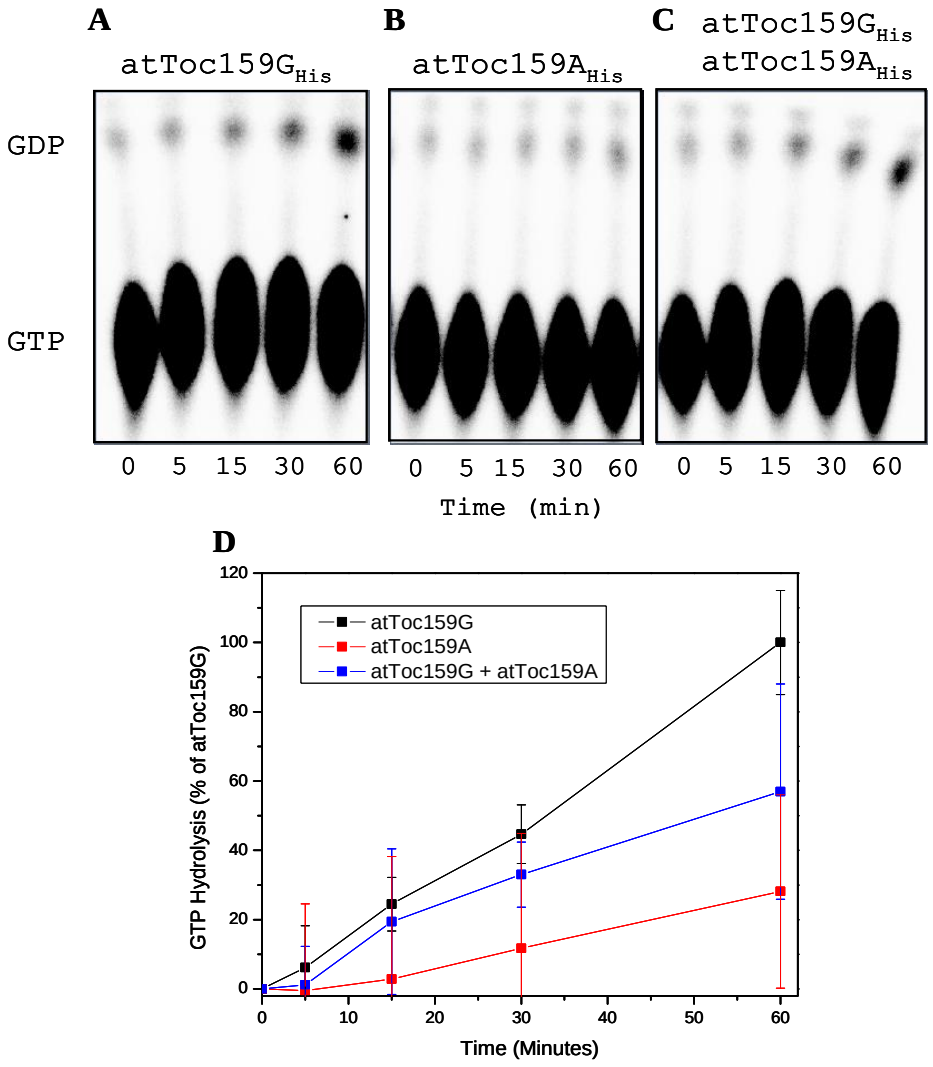
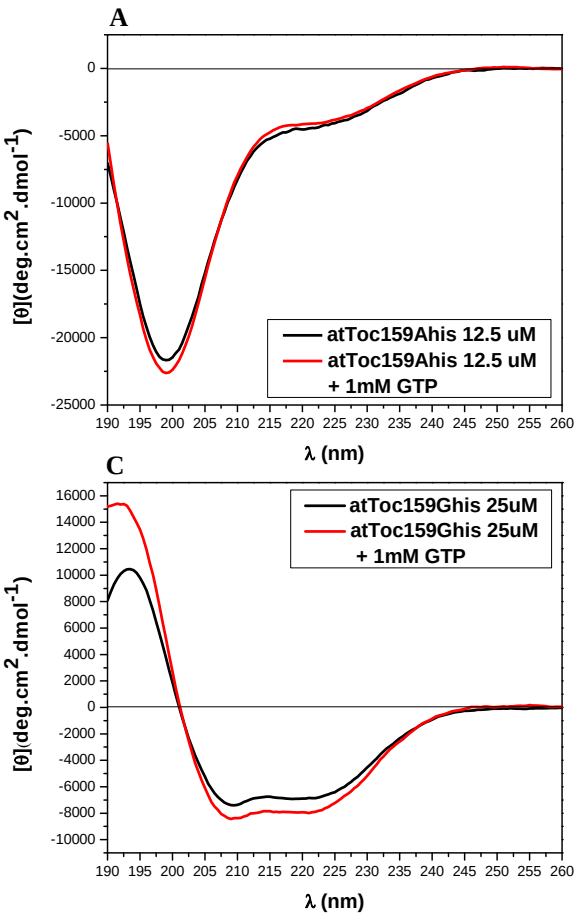


Figure 9. The effect of A-domain on the GTPase activity of the G-domain. Hydrolysis of α-[^{32}P]-GTP by atToc159G$_{His}$ (A), atToc159A$_{His}$ (B), and a mixture of both proteins (C) was measured over time and nucleotide species were resolved by TLC. TLC plates were analyzed by phosphor-imaging (A, B, C) and a ratio of GDP to GTP was used to calculate hydrolysis for atToc159G$_{His}$. The activities of the other samples were expressed as a percentage of the hydrolysis by atToc159G$_{His}$ at 60 min (D). Data collected from two independent experiments.



Structure	atToc159A$_{His}$ (12.5 µM)	
	- GTP	+ GTP
α-Helical	4 %	5 %
β-Strand	7 %	4 %
β-Turn	6 %	5 %
Random Coil Disordered	81 %	85 %

Structure	atToc159G$_{His}$ (25 µM)	
	- GTP	+ GTP
α-Helical	12 %	24 %
β-Strand	24 %	18 %
β-Turn	18 %	21 %
Random Coil Disordered	36 %	27 %

Figure 10. CD spectra of atToc159A$_{His}$ (A) and atToc159G$_{His}$ (C) at 12.5 µM and 25 µM, respectively, from 190-260 nm measured with (red trace) and without (black trace) 1 mM GTP. The spectra were deconvoluted using Dichroweb (CDSSTER, ref set #7) and the proportions of various secondary structure elements were estimated for each (Figure 10, B, D). Data collected from four independent experiments.

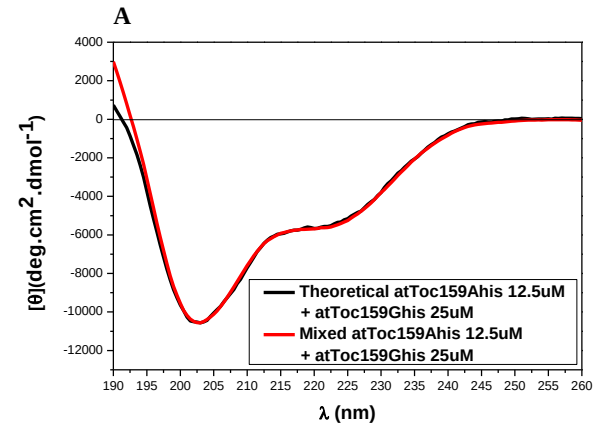
This increase in α-helix content is reflected in the secondary structure estimation (Figure 10, D) which shows increase in α-helix and decrease in disordered structures.

3.6 Changes in secondary structure induced by the interaction between the A- and G-domains of atToc159 are not measurable by circular dichroism spectropolarimetry

The CD spectra of atToc159A_{His} and atToc159G_{His} were measured as individual proteins and as mixtures at ratios of 1:2 (Figure 11a) and 1:1 (Figure 11b). Comparison of added individual CD spectra of atToc159 A- and G-domains (black trace) and the mixture of these two proteins (red trace) could reveal whether or not any significant change in secondary structure occurred upon the interaction between these two protein domains. The proteins were measured in the presence and absence of GTP (1 mM) to determine whether or not the nucleotide-state of atToc159G_{His} played a role in the interaction. The addition of individual spectra and the observed spectrum from the mixed samples were not noticeably different regardless of the presence of GTP. This is reflected in the marginal changes in secondary structure as estimated by deconvolution (Dichroweb; CDSSTER, ref set #7) (Figure 11a, 11b - B, D), indicating that neither the A-domain nor the G-domain underwent a significant structural change when the two proteins were mixed at the concentrations tested.

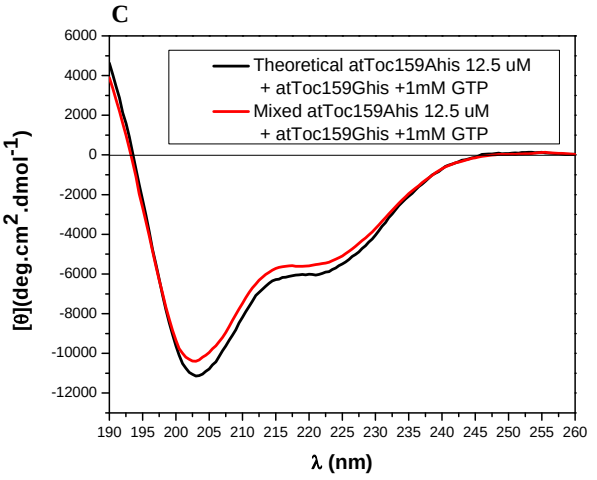
3.7 The interaction between the Toc159 A-domain and actin filaments.

To directly test whether the A-domain of atToc159 plays a role in the observed interaction between the pre-protein receptor and actin filaments (Jouhet & Gray



B

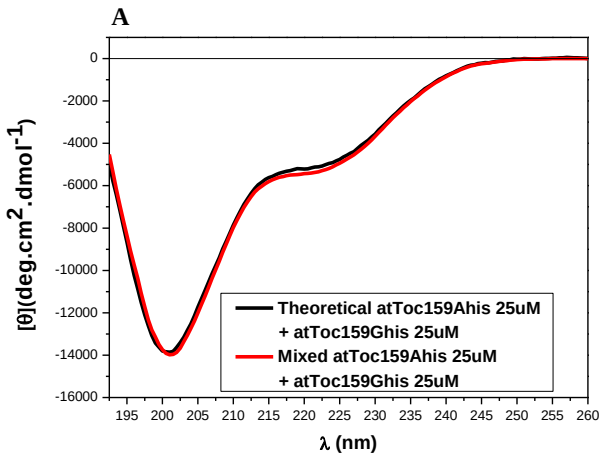
Structure	- GTP	
	Theoretical	Mixture
α-Helical	14 %	11 %
β-Strand	15 %	15 %
β-Turn	14 %	12 %
Random Coil Disordered	56 %	61 %



D

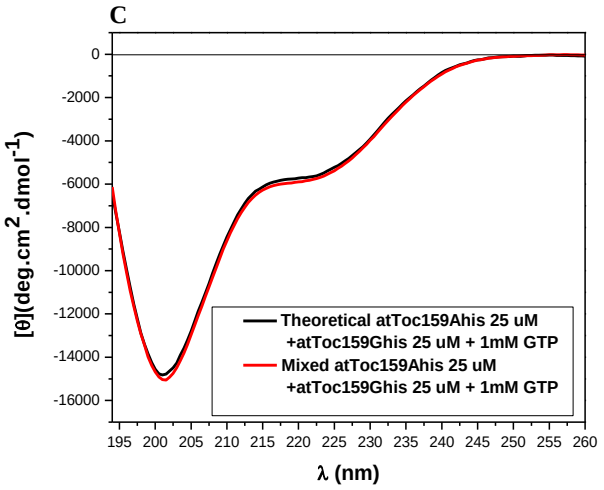
Structure	+ GTP (1 mM)	
	Theoretical	Mixture
α-Helical	16 %	9 %
β-Strand	17 %	16 %
β-Turn	14 %	15 %
Random Coil Disordered	52 %	59 %

Figure 11a. CD spectra of atToc159A$_{His}$ and atToc159G$_{His}$ at 1:2 ratio measured with (Figure 11a, A) and without (Figure 11a, C) GTP (1 mM). The addition of individual spectra (theoretical - black trace) represents a mixture void of changes in secondary structure to be compared to the spectra of the mixture of the two protein domains (mixture - red trace). Secondary structure was estimated by deconvolution (Dichroweb; CDSSTER, ref set #7) for both the theoretical and mixed spectra (Figure 11a, B, D). Data collected from four independent experiments.



B

Structure	- GTP	
	Theoretical	Mixture
α-Helical	7 %	7 %
β-Strand	11 %	12 %
β-Turn	11 %	11 %
Random Coil Disordered	70 %	68 %



D

Structure	+ GTP	
	Theoretical	Mixture
α-Helical	8 %	7 %
β-Strand	10 %	12 %
β-Turn	10 %	10 %
Random Coil Disordered	70 %	70 %

Figure 11b. CD spectra of atToc159A$_{His}$ and atToc159G$_{His}$ at 1:1 ratio measured with (Figure 11b, A) and without (C) GTP (1 mM). The addition of individual spectra (theoretical - black trace) represents a mixture void of changes in secondary structure to be compared to the spectra of the mixture of the two protein domains (mixture - red trace). Secondary structure was estimated by deconvolution (Dichroweb; CDSSTER, ref set #7) for both the theoretical and mixed spectra (Figure 11b, B, D). Data collected from four independent experiments.

2009), co-sedimentation assays with recombinant atToc159A and actin filaments were conducted. When 100 pmol of F-actin on its own was sedimented, greater than 90% of actin was found in the pellet (Figure 12, lane 2) and only trace amounts were observed in the soluble fraction (Figure 12, lane 3). Interestingly, the sedimentation of 100 pmol of F-actin in the presence of 50 pmol of atToc159A$_{His}$ caused a shift in the distribution of F-actin, such that it was approximately equally distributed between the pellet and supernatant. On the other hand, only 10% of the atToc159A was found in the pellet fraction, whereas approximately 90% remained in the supernatant (Figure 12, lanes 4 and 5).

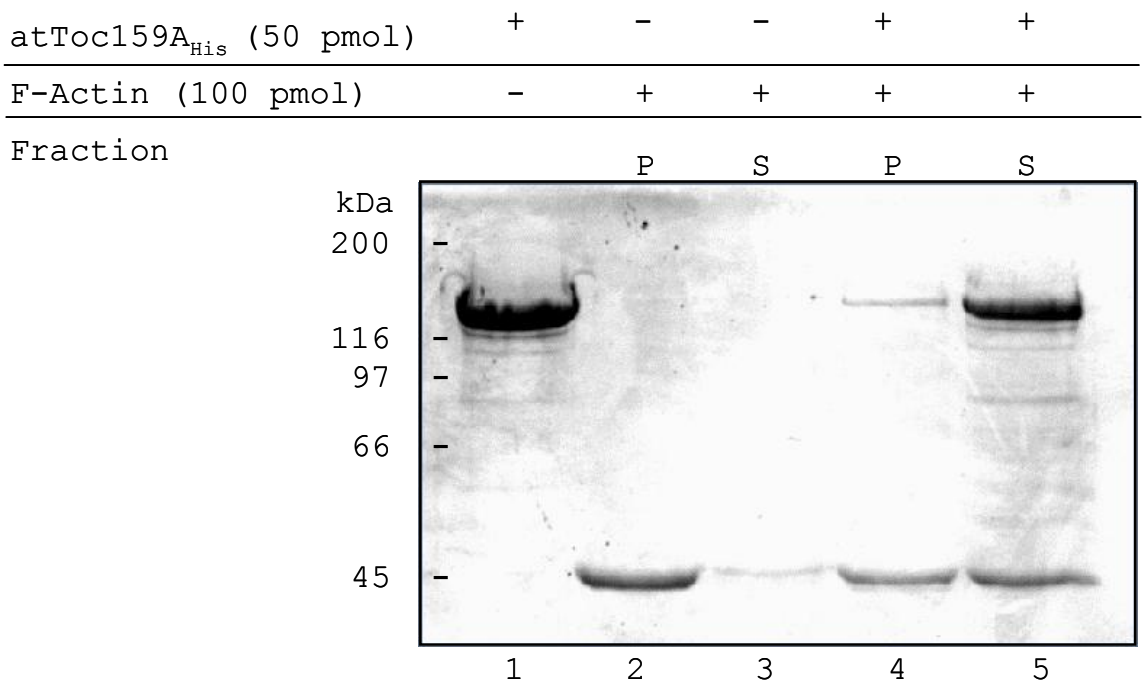


Figure 12. atToc159A-actin co-sedimentation assay. atToc159A was mixed with F-actin, the mixture was centrifuged to pellet F-actin, and pellet and supernatant fractions were analyzed using SDS-PAGE stained with Coomassie Blue. Lane 1 was loaded with 50 pmol of atToc159A$_{His}$, lanes 2 and 3 were loaded with the pellet and supernatant, respectively, from 100 pmol of sedimented F-actin, lanes 4 and 5 were loaded with the pellet and supernatant, respectively, from 50 pmol of atToc159A$_{His}$ co-sedimented with 100 pmol F-actin.

4. Discussion

In this study, a number of *in vitro* techniques were used to: 1) demonstrate that the A-domains of the Toc159 homologues interact specifically with their respective G-domains; and 2) characterize the functional role that this interaction plays in chloroplast pre-protein import. The initial observation that the free A-domain associates with the chloroplast was from the study by Agne et al. (2010) and an observation by Inoue & Schnell (unpublished, personal communication). The limited proteolysis of wild-type chloroplasts, repeated during the course of the current study with thrombin, shows that the A-domain of atToc159 remains associated with the chloroplast when proteolytically separated from the G-domain (Figure 6). However, *in vitro* translated atToc159A was not observed to target to atToc159GM$_{His}$ /*ppi2* (atToc159A-deficient) chloroplasts, and targeting to wild-type chloroplasts was minimal (Figure 7). To test the hypothesis that the atToc159 A-domain remained associated with chloroplasts through an interaction with the atToc159 GTPase domain, solid-phase binding assays were conducted. These assays demonstrate that the A-domains of atToc159 and atToc132 interact in a dose-dependent manner with their respective G-domains without cross-reactivity with the GTPase domains of the other Toc159 homologue nor with the Toc34 homologues (Figure 8). To functionally characterize this specific interaction, the GTPase activity of atToc159G was analyzed in the presence and absence of atToc159A. While the use of the TLC-based assay lacked sensitivity and the results were difficult to reproduce with small measures of error, the data suggest that the A-domain has an inhibitory effect on GTPase activity of atToc159G (Figure 9). In an attempt to structurally characterize the interaction

between the A- and G-domains of atToc159, circular dichroism spectropolarimetry was employed. The A-domain did not show any structural changes upon the addition of GTP, but the G-domain exhibited an increase in α-helical character (Figure 10). When the A-domain and G-domain were mixed in molar ratios of 1:1 in the absence and presence of GTP no changes in secondary structure were detected (Figures 11b). However, when the A-domain and G-domain were mixed in molar ratios of 1:2, a slight decrease in secondary structure was observed, most notably in the presence of GTP. While these experiments do not demonstrate that a significant conformational change accompanies the interaction between these two protein domains at the concentrations and ratios measured, it is possible that a more substantial loss of structure could be observed by CD under alternative conditions.

4.1 The A-domain interacts specifically with the G-domain of Toc159

The association of a portion of the A-domain with the chloroplast membrane following proteolytic cleavage with thrombin corroborates a study by Agne et al. (2010), in which TAP (tandem affinity purification)-tagged atToc159A was expressed in wild-type plants and found in both soluble and chloroplast fractions. The plant-derived atToc159A was also found to be heavily phosphorylated (Agne et al. 2010). *In vitro* translated [^{35}S]-atToc159A produced using a rabbit reticulocyte lysate system is not post-translationally modified in the same way as atToc159 would be in the cytoplasm of the plant cell, and at the surface of the chloroplast. May & Soll (2000) observed that rabbit reticulocyte lysate derived protein did not behave the same as wheat germ derived protein suggesting that *in vitro* translated products do not always

behave the same as *in vivo* derived proteins. It is possible that a lack of post-translational modification contributed to the differences in A-domain localization observed between the thrombin treated chloroplasts (Figure 6) and the *in vitro* targeting of [^{35}S]-atToc159A to isolated chloroplasts, which exhibited extremely low levels of targeting (Figure 7). This low-level targeting could also be attributed to the nature of the interaction and the targeting assay itself. Specifically, protein-protein interactions involving IUPs such as the A-domain are characteristically low affinity and often transient in nature (Dyson et al. 2005). As such, the re-isolation of chloroplasts in percoll following the targeting phase of the assay may have resulted in dissociation of the A-domain from the G-domain at the surface of the chloroplast.

The solid-phase binding assays demonstrated that the G-domain of Toc159 can bind the A-domain suggesting that it could be the site of interaction responsible for the localization of free A-domain at the chloroplast outer envelope membrane (Figure 8). The A-domains of atToc159 and atToc132 bound specifically to their respective G-domains, but not those of the other Toc159 homologue, or with that of atToc33 or atToc34. The level of binding of [^{35}S]-atToc159A to immobilized atToc159G and isolated chloroplasts is very different (~15% vs. <1%, respectively). This variability can be attributed to the significant difference in available G-domain in the two assays. It is possible however, that the post-translational phosphorylation of *in vivo* derived A-domain is not essential for the interaction between the A-domain and G-domain of Toc159, but it may affect the interaction in the more complex environment at the chloroplast outer envelope membrane.

In an effort to fully characterize the interaction between the A-domain and G-domain of Toc159, biophysical and biochemical analyses of the interaction were conducted. Circular dichroism spectropolarimetry was used to examine the secondary structure of the A- and G-domain of atToc159 as single protein samples and in combination to examine the structure associated with the interaction (Figure 11). The potential biochemical function of the interaction was assessed by examining the GTP hydrolysis of the G-domain of atToc159 in the absence and presence of the A-domain (Figure 9).

The A-domain produces a high-intensity CD spectrum that, when subject to deconvolution, demonstrates that the domain is dominated by random coil or disordered structure (Figure 10, A, B; Richardson et al. 2009). Conversely, the G-domain possesses significant α-helical content, more so in the presence of GTP (Figure 10). It is possible for intrinsically unstructured proteins to adopt secondary structure elements as a result of protein-protein interactions based on the conformation of the ligand (Dyson et al. 2002). However, a large conformational change in the A-domain was not observed upon mixing with the G-domain (Figure 11). These data suggest that the interaction between these two protein domains is not accompanied by a large structural change in the A-domain at the concentrations and ratios tested. *In vivo*, the effective concentration of proteins close together at a membrane surface is higher than those in the soluble phase, especially when those proteins may be covalently linked. The large size of the A-domain and the extremely high number of solvent-exposed residues suggests that the interaction with the G-domain may involve only a small portion of the A-domain. If this were the case, a

gain in structure of the interacting region would not greatly affect the overall character of the A-domain and would be undetectable by CD. It would also have implications in the proposed function of the A-domain. While it has been demonstrated that there is no direct interaction between pre-proteins and the A-domain *in vitro*, it is possible that the interaction between the A-domain and G-domain *in vivo* is involved in mediating an interaction with pre-proteins.

Another possible function of the A-domain – G-domain interaction is the regulation of GTP hydrolysis by the G-domain. Previously, the G-domain of atToc159 has been shown to bind pre-proteins in a GTP-bound state (Wang et al. 2008). The observed inhibition of atToc159 GTPase activity elicited by the A-domain (Figure 9) suggests that it might play a role in the nucleotide-dependent pre-protein recognition by the G-domain. This is corroborated by reports of IUP domains interacting with other enzymatic domains of the same protein to modulate activity (Jonker et al. 2006). Recent experiments examining the dimerization and GTP hydrolysis of atToc33 have demonstrated that dimerization of this Toc component acts to inhibit nucleotide exchange (Oreb et al. 2011). The dimerization of the G-domain of atToc159 has not been studied as extensively as that of atToc33, but increasing concentrations of atToc159G, as monitored by CD, suggest that there may be some association at higher concentrations corresponding to dimer formation (see Appendix III). If dimerization does occur, it is possible that the A-domain is involved in a mechanism similar to that of pre-proteins in their interaction with atToc33. The interaction between the G-domain in its GTP bound state and pre-proteins may have some influence on the G-domain – A-domain interaction that results in dissociation of

the A-domain followed by GTP hydrolysis. It is also possible that the A-domain is involved in the discretion that full length atToc159 demonstrates in pre-protein binding by influencing the transit peptide binding site.

An alternate putative function of the A-domain, based on work by Jouhet & Gray (2009), is the binding of actin cytoskeleton to assist in chloroplast mobility. While other chloroplastic proteins such as CHUP1 (chloroplast unusual positioning 1) and the gene product knocked out in *pmi1* (plastid movement impaired 1) interact with actin and are important for light-dependent chloroplast movement, the complete mechanism of intracellular chloroplast mobility is undetermined (Oikawa et al. 2003; DeBlasio et al, 2005). Jouhet et al. (2009) observed that recombinant full length atToc159 could be purified from *E. coli* lysate by co-sedimentation with actin, but purified atToc159G and atToc33G did not bind actin. In actin co-sedimentation assays, only ~10% of purified atToc159A in the reaction was found to sediment with actin (Figure 12). However, instead of finding ~90% of the actin in the pellet fraction as observed in the control, approximately 50% remained in the soluble fraction with the soluble atToc159A (Figure 12). Monomeric (G)-actin requires the presence of divalent cations, primarily Mg^{2+}, potassium (K^{+}), and ATP to form filaments. ATP-bound monomers associate with existing filaments, comprised largely of ADP-bound actin monomers, and the hydrolysis of ATP strengthens the interaction between the monomer and the filament (Dos Remedos et al. 2003). The assembly and dissociation of actin filaments is mediated by a large number of actin-binding proteins (ABP). We conclude from these data that the heavily charged A-domain could be interacting with F-actin filaments thereby altering their solubility or acting as an ABP

that causes the actin filaments to dissociate to a monomeric state. These are inconclusive preliminary results but are interesting nonetheless. Further studies concerning this interaction to clarify the nature of the A-domain – actin interaction are required.

4.2 Future directions for elucidating the role of the atToc159 A-domain – G-domain interaction.

The specific interaction between the A- and G-domains of atToc159 and the observed inhibition of GTP hydrolysis suggests that the A-domain is involved in the early stages of pre-protein recognition. To strengthen the conclusions drawn from this work and further characterize the role of the A-domain in pre-protein recognition by structurally distinct Toc complexes, further experimentation with higher levels of complexity is required. The low affinity, transient interactions that generally characterize protein-protein interactions involving IUPs – and the type of interaction required to ensure efficient translocation by protein translocons - pose problems when using traditional binding and targeting assays to characterize them, which often rely on stronger, more stable interactions. The use of chemical cross-linking in chloroplast protein targeting assays, similar to that used in the study by Smith et al (2004), would potentially eliminate the problem posed by transient interactions, as it would prevent the dissociation of any A-domain targeted to chloroplasts and allow for the identification of other proteins, such as pre-proteins, involved in the interaction.

It is unlikely that the recognition and binding of pre-proteins at the chloroplast is as simple as a single protein-protein interaction, given the complexity of other

interactions in chloroplast pre-protein import and the general Toc complex structure. Introducing higher levels of complexity into the assays used in the current study would shed light on the nature of the interaction observed between the A-domain and G-domain of atToc159. By including pre-proteins together with the A-domain in the solid phase binding assays with immobilized G-domain, information about how the A-domain influences pre-protein binding could be gained.

This same principle should be applied to the GTP hydrolysis assays. Examining GTPase activity of the G-domain in the presence of pre-proteins and the A-domain would provide further insight into the function of the A-domain's inhibition of G-domain GTPase activity. Along with the GTP hydrolysis, the dissociation of GDP using fluorescently labeled mant-GDP in the presence of pre-proteins and the A-domain should be examined, similar to the analysis of atToc33 by Oreb et al. (2011). The nucleotide binding state of the G-domain during interactions with the A-domain would provide further insight into how the A-domain moderates pre-protein binding.

Another piece of information missing from the pre-protein recognition story is the nature and potential importance of atToc159 G-domain dimerization. The CD data presented in Appendix II is incomplete and requires additional work. Repeating those experiments over a larger range of concentrations would clarify the observed trend. Also measuring near-UV CD of the same samples would clarify the far-UV data, and whether or not proteins were associating, by giving information about the environment of aromatic residues. Another, more direct approach that could be used to examine G-domain dimerization is blue-native PAGE. If the G-domain of

atToc159 does form dimers, experiments similar to those conducted by Oreb et al (2011) could be conducted to examine how the dimer influences GTP hydrolysis and how the A-domain and/or pre-protein influence the dimer.

In the study by Agne et al (2010), the A-domain was shown to be heavily phosphorylated *in vivo*. They determined this by using a TAP-tagged atToc159A product purified from plant tissue. The role of phosphorylation of the A-domain has not been elucidated as the inhibition of phosphorylation did not affect the import of prSSU (Agne et al. 2010). While the rate of import was not affected by phosphorylation, the post-translational modification of the A-domain may have implications in the interaction with the G-domain or the ability of distinct Toc complexes to differentiate photosynthetic pre-proteins from non-photosynthetic pre-proteins. The use of plant purified A-domain rather than *E. coli*-produced protein, in the chloroplast binding and nucleotide binding and hydrolysis assays described in the current study might yield different information concerning the effect of post-translational modification on the binding of pre-proteins. Future experiments could substitute plant-produced and purified proteins in these assays.

While collecting data from *in vitro* experiments, like those outlined in this section, will be important, more *in vivo* approaches are also needed. The immense complexity of the chloroplast protein translocon, with a large number of essential and regulatory components, requires analysis as a whole as well as in parts. Once the role of the A-domain interactions have been more clearly defined, movement back to the chloroplast and back to the plant will be crucial.

4.3 Integration of multiple approaches with a common goal

Chloroplast protein import plays a crucial role in the growth and development of leafy green plant tissue. Protein trafficking is an important part of cellular function in all eukaryotes, and our understanding of systems like the chloroplast protein translocon shed light on the function and regulation of similar processes. The study of protein import into chloroplasts is important for characterizing intracellular protein import in general. In addition, plants are essential to the world-wide ecosystem and provide us with a vital food source. The development of new crop-growing strategies and plant manipulation is important, given the rate at which food demands are increasing, and because chloroplasts play such an instrumental role in plant development, their understanding is essential. This requires the collaboration of work from a number of different fields of expertise. The current knowledge concerning the chloroplast protein import pathway has come from a diverse and vast body of research involving evolutionary (Dyall et al. 2004) and molecular biology (Ivanova et al. 2004; Smith et al. 2004; Inoue et al 2010), biochemical (Richardson et al. 2009) and *in vivo* (Bauer et al. 2000; Kubis et al. 2003) based approaches. The identification of various Toc-Tic constituents and the characterization of their function have come from the development and use of knockout mutants, elegant biochemical and biophysical assays, the use of recombinant DNA technology, and large scale genetic and proteomic studies. The integration of a vast amount of diverse research has led to our current understanding of chloroplast protein import and how it fits into the larger picture of plant development. The continued development of new

approaches and collaboration by a number of research groups is important for the future as our understanding of plant biology grows.

5. Conclusions

It was hypothesized that the Acidic-domain of atToc159 interacts with the GTPase-domain of atToc159 to confer the functional specificity observed between structurally distinct Toc complexes by regulating the Toc159 GTP hydrolysis involved in pre-protein import. The current study has demonstrated a specific, dose-dependent interaction between the A-domain and G-domain of atToc159 that is likely responsible for the observed association of the A-domain with the outer membrane of the chloroplast and appears to inhibit the GTP hydrolysis activity of the G-domain. The precise implications this interaction has for chloroplast protein import is unknown, but it is reasonable to speculate that it is involved in the initial steps of pre-protein recognition by the Toc complex, and contributes to the observed specificity between functionally distinct complexes.

The current work needs to be expanded to involve higher levels of complexity, including pre-proteins and other Toc complex constituents in experiments previously conducted to determine the role of the A-domain - G-domain interaction in pre-protein binding.

6. Work Cited

Agne, B., Andres, C., Montandon, C., Christ, B., Ertan, A., Jung, F., Infanger, S., Bischof, S., Baginsky, S., Kessler, F. The acidic A-domain of *Arabidopsis* Toc159 occurs as a hyperphosphorylated protein. *Plant Physiology*. **153** 1016-1030. (2010)

Agne, B., Kessler, F. Protein transport in organelles: the Toc complex way of pre-protein import. *FEBS Journal*. Author's Compilation, 1-10 (2009)

Armbruster, U., Hertle, A., Makarenko, E., Zuhlke, J., Pribil, M., Dietzmann, A., Schliebner, I., Aseevna, E., Fenino, E., Scharfenberg, M., Voigt, C., Leister, D. Chloroplast proteins without cleavable transit peptides: rare exceptions or a major constituent of the chloroplast proteome? *Molecular Plant*. **2** 1325-1335 (2009)

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

Aronsson, H., Combe, J., Patel, R., Agne, B., Martin, M., Kessler, F., Jarvis, P. 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.* 1-15 (2010)

Bae, W., Lee, Y.J., Kim, D.H., Lee, J., Kim, S., Sohn, E.J., Hwang, I. AKR2A-mediated import of chloroplast out membrane proteins is essential for chloroplast biogenesis. *Nature Cell Biology*. **10** 220-227 (2008)

Baldwin, A.J., Inoue, K. The most C-terminal tri-glycine segment within the polyglycine stretch of the pea Toc75 transit peptide plays a critical role for targeting the protein to the chloroplast outer envelope membrane. *FEBS*. **273** 1547-1555 (2006)

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

Bannai, H., Tamada, Y., Maruyama, O., Nakai, K., Miyano, S. Extensive feature detection of N-terminal protein sorting signals. *Bioinformatics*. **18** 298-305 (2002)

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

Becker, T., Hritz, J., Vogel, M., Caliebe, A., Bukau, B., Soll, J., Schleiff, E. Toc12, a novel subunit of the intermembrane space pre-protein translocon of chloroplasts. *Molecular Biology of the Cell*. **15** 5130-5144 (2004)

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

Bischof, S., Baerenfaller, K., Wildhaber, T., Troesch, R., Vidi, P.A., Roschitzki, B., Hirsch-Hoffmann, M., Hennig, L., Kessler, F., Gruissem, W., Baginsky, A. Plastid proteome assembly without Toc159: Photosynthetic protein import and accumulation f *N*-Acetylated plastid precursor proteins. *The Plant Cell*. **23** 3911-3928 (2011)

Brock, I., Hazell, L., Muchl, D., Nielsen, V.S., Moller, B.L., Herrmann, R.G., Klosgen, R.B., Robinson, C. *Plant Molecular Biology*. **23** 717-725 (1993)

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

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

Chen, X., Smith, M.D., Fitzpatrick, L., Schnell, D.J. *In vivo* analysis of the role of Tic20 in protein import into chloroplasts. *The Plant Cell*. **14** 641-654 (2002)

Chen, Yi. The role of the acidic domain of the Toc159 protein import receptor family in Toc complex assembly in *Arabidopsis thaliana*. *Wilfrid Laurier University Graduate Thesis*. (2011)

Chiu, C.C., Chen, L.J., Li, H.M. Pea chloroplast DnaJ-J8 and Toc12 are encoded by the same gene and localized in the stroma. *Plant Physiology*. **154** 1172-1182 (2010)

Chou M.L., Chu, C.C., Chen, L.J., Akita, M., Li, H.M. Stimulation of transit-peptide release and ATP hydrolysis by a co-chaperone during protein import into chloroplasts. *Journal of Cell Biology*. **175** 893-900 (2006)

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

DeBlasio, S.L., Luesse, D.L., Hangarter, R.P. A plant-specific protein essential for blue-light-induced chloroplast movements. *Plant Physiology*. **139** 101-114 (2005)

Dhanoa, P.K., Richardson, L.G.L., Smith, M.D., Gidda, S.K., Henderson, M.P.A., Andrews, D.W., Mullen, R.T. Distinct pathways mediate the sorting of tail-anchored proteins to the plastid outer envelope. *Public Library of Science*. **5** (4) 1-18 (2010)

Dos Remedos, C.G., Chhabra, D., Kekic, M., Dedova, I.V., Tsubakihara, M., Berry, D.A., Nosworthy, N.J. Actin binding proteins: Regulation of cytoskeletal microfilaments. *Physiology Review*. **83** 433-473 (2003)

Dyall, S.D., Brown, M.T., and Johnson, P.J. Ancient invasions: From endosymbionts to organelles. *Science.* **304** 253-257 (2004)

Dyson, J. H., Wright, P. E. Coupling of folding and binding for unstructured proteins. *Current Opinion in Structural Biology*. **12** 54-60 (2002)

Dyson, H.J., and Wright, P.E. Intrinsically unstructured proteins and their functions. *Nature Review Molecular Cell Biology.* **6** 197-208 (2005)

Eckart, K., Eichacker, L., Sohrt, K., Schleiff, E., Heins, L., Soll, J. A Toc75-like protein import channel is abundant in chloroplasts. *EMBO*. **3** 557-562 (2002)

Emanuelsson, O., Nielsen, H., and von Heijne, G. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their sequences. *Protein Science.* **8** 978-984 (2005)

Fink, A.L. Natively unfolded proteins. *Current Opinion in Structural Biology*. **15** 35-41 (2005)

Firlej-Kwoka, K., Strittmatter, P., Soll, J., Bolter, B. Import of pre-proteins into the chloroplast inner envelope membrane. *Plant Molecular Biology*. **68** 505-519 (2008)

Froehlich, J.E., Keegstra, K. The role of the transmembrane domain in determining the targeting of membrane proteins to either the inner envelope or thylakoid membrane. *The Plant Journal*. **68** 844-856 (2011)

Fuxreiter, M., Tompa, P., Simon, I., Uversky, V.N., Hansen, J.C., Asturias, F.J. Malleable machines take shape in eukaryotic transcriptional regulation. *Nature Chemical Biology*. **4** 728-737 (2008)

Galea, C.A., Wang, Y., Sivakolundu, S.G., Kriwacki, R.W. Regulation of cell division by intrinsically unstructured proteins: Intrinsic flexibility, modularity, and signaling conduits. *Biochemistry*. **47** 7598-7609 (2008)

Greenfield, N.J. Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols.* **1** 2876-2890 (2007)

Gunasekaran, K., Tsai, C.J., Kumar, S., Zanuy, D., Nussinov, R. Extended disordered proteins: targeting function with less scaffold. *TRENDS in Biochemical Sciences.* **28** 81-85 (2003)

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

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

Hirohasi, T., Hase, T., Nakai, M. Maize non-photosynthetic ferredoxin precursor is miss-sorted to the intermembrane space of chloroplasts in the presence of light. *Plant Physiology*. **125** 2154-2163 (2001)

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

Hsu, S.C., Inoue, K. Two evolutionarily conserved essential β-barrel proteins in the chloroplast outer envelope membrane. *BioScience Trends*. **3** 168-178 (2009)

Hsu, S.C., Nafati, M., Inoue, K. OEP80, an essential protein paralogous to the chloroplast protein translocation channel Toc75, exists as a 70-kD protein in the *Arabidopsis thaliana* chloroplast outer envelope. *Plant Molecular Biology*. **78** 147-158 (2012)

Inaba, T., Li, M., Alvares-Huerta, M., Kessler, F., Schnell, D.J. atTic110 functions as a scaffold for coordinating the stromal events of protein import into chloroplasts. *Journal of Biological Chemistry*. **278** 38617-38627 (2003)

Inaba, T., Schnell, D.J. Protein trafficking to plastids: one theme, many variations. *Biochemical Journal*. **413** 15-28 (2008)

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

Inoue, K., Baldwin, A.J., Shipman, R.L., Matsui, K., Theg, S.M., Ohme-Takagi, M. Complete maturation of the plastid protein translocon channel requires a type I signal peptidase. *Journal of Cell Biology*. **171** 425-430 (2005)

Inoue, K., Demel, R., de Kruijff, B., Keegstra, K. The N-terminal portion of the pre-Toc75 transit peptide interacts with membrane lipids and inhibits binding and import of precursor proteins into isolated chloroplasts. *European Journal of Biochemistry*. **268** 4036-4043 (2001)

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

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

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

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

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

Jelic, M., Svenshnikova, N., Motzkus, M., Horth, P., Soll, J., Schleiff, E. The chloroplast import receptor Toc34 functions as a pre-protein-regulated GTPase. *Journal of Biological Chemistry*. **383** 1875-1883 (2002)

Jelic, M., Soll, J., Schleiff, E. Two Toc34 with different properties. *Biochemistry*. **42** 5906-5916 (2003)

Jonker, H.R.A., Wechselberger, R.W., Boelens, R., Kaptein, R., Folkers, G.E. The intrinsically unstructured domain of PC4 modulates the activity of the structured core through inter- and intramolecular interactions. *Biochemistry*. **45** 5067-5081 (2006)

Jouhet, J., Gray, J.C. Interaction of actin and the chloroplast protein import apparatus. *Journal of Biological Chemistry*. **284** 19132-19141 (2009)

Kaneko, T., Sato, S., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoti, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., Tabata, S. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. Strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Research.* **3** 109-136 (1996)

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

Keeling, P.J. Diversity and evolutionary history of plastids and their hosts. *American Journal of Botany*. **91** 1481-1493 (2004)

Kelly, S.M., Price, N.C. The use of circular dichroism in the investigation of protein structure and function. *Current Protein and Peptide Science* 349-384 (2000)

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

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

Kim, D.H., Xu, Z-Y., Na, Y.J., Yoo, Y-J., Lee, J., Sohn, E-J., Hwang, I. Small heat shock protein Hsp17.8 functions as an AKR2A cofactor in the targeting of chloroplast outer membrane proteins in *Arabidopsis*. *Plant Physiology*. **157** 132-146 (2011)

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

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

Kouranov, A., Wang, H., Schnell, D.J. Tic22 is targeted to the intermembrane space of chloroplasts by a novel pathway. *Journal of Biological Chemistry*. **274** 25181-25186 (1999)

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

Kovacs-Bogdan, E., Benz, J.P., Soll, J., Bolter, B. Tic20 forms a channel independent of Tic110 in chloroplasts. *BMC Plant Biology*. **11**:133 1-16 (2011)

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

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

Lee, D.W., Lee, S., Oh, Y.J., Hwang, I. Multiple sequence motifs in the RbcS transit peptide independently contribute to Toc159-dependent import of proteins into chloroplasts. *Plant Physiology*. **151** 129-141 (2009)

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

Li, H., Chiu, C. Protein transport into chloroplasts. *Annual Review of Plant Biology*. **61** 157-180 (2010)

Li, M., Schnell, D.J. Reconstitution of protein targeting to the inner envelope membrane of chloroplasts. *Journal of Cell Biology*. **175** 249-259 (2006)

Lobley, A., Whitmore, L. and Wallace, B.A. DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics.* **18** 211-212 (2002)

Lubeck, J., Heins, L., Soll, J. A nuclear-coded chloroplastic inner envelope membrane protein uses a soluble sorting intermediate upon import into the organelle. *Journal of Cell Biology*. **137** 1279-1286 (1997)

Lubeck, J., Soll, J., Akita, M., Nielsen, E., Keegstra, K. Topology of IEP110, a component of the chloroplastic protein import machinery present in the inner envelope membrane. *EMBO Journal*. **15** 4230-4238 (1996)

Martin, W., Herrmann, R.G. Gene transfer from organelles to the nucleus: how much, what happens and why? *Plant Physiology.* **188** 9-17 (1998)

Mathews, C.K. and Van Holde, K.E. Biochemistry 2nd Ed. Benjamin-Cummings Publishing. 204-210 (1996)

May, T., and Soll, J. 14-3-3 proteins for a guidance complex with chloroplast precursor proteins in plants. *The Plant Cell*. **12** 53-64 (2000)

Miras, S., Salvi, D.S., Ferro, M., Grunwald, D., Garin, J., Joyard, J., Rolland, N. Non-canonical transit peptide for import into the chloroplast. *Journal of Biological Chemistry*. **277** 47770-47778 (2002)

Miras, S., Salvi, D., Piette, L., Seigneurin-Berny, D., Grunwald, D., Reinbothe, C., Joyard, J., Reinbothe, S., and Rolland, N. Toc159-and Toc75-independent import of a transit sequence-less precursor into the inner envelope of chloroplasts. *Journal of Biological Chemistry*. **282** 29482–29492 (2007)

Nada,A., Soll, J. Inner envelope protein 32 is imported into chloroplasts by a novel pathway. *Journal of Cell Science*. **117** 3975-3982 (2004)

Nielsen, E., Akita, M., Davlia-Aponte, J., Keegstra, K. Stable association of the chloroplastic precursors with protein translocation complexes that contain proteins from both envelope membranes and a stromal Hsp100 molecular chaperone. *EMBO Journal*. **16** 935-946 (1997)

Oikawa, K., Kasahara, <., Kiyosue, T., Suetsugu, N., Takahashi, F., Kanegae, T., Niwa, Y., Kadota, A., Wada, M. Chloroplast Unusual Positioning1 is essential for proper chloroplast positioning. *The Plant Cell*. **15** 2805-2815 (2003)

Oreb, M., Hofle, A., Koenig, P., Sommer, M.S., Sinning, I., Wang, F., Tews, I., Schnell, D.J., Schleiff, E. Substrate binding disrupts dimerization and induces nucleotide exchange of the chloroplast GTPase atToc33. *Biochemical Journal*. **436** 313-319 (2011)

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

Putnam, C. The Scripps Research Institute, U.S.A..
<http://www.scripps.edu/~cdputnam/protcalc.html>

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

Raven, J.A., Allen, J.F. Genomics and chloroplast evolution: what did cyanobacteria do for plants? *Genome Biology*. **4** 209 (2003)

Reddick, L.E., Vaughn, M.D., Wright, S.J., Campbell, I.M., Bruce, B.D. *In vitro* comparative kinetic analysis of the chloroplast Toc GTPases. *Journal of Biological Chemistry*. **282** 11410-11426 (2007)

Reumann, S., Inoue, K., Keegstra, K. Evolution of the general protein import pathway of plastids. *Molecular Membrane Biology*. **22** 73-86 (2005)

Richardson, L.G.L, Jelokhani-Niaraki, M., Smith, M.D. The acidic domains of the Toc159 chloroplast pre-protein receptor family are intrinsically disordered protein domains. *BMC*. **10**:35 1-8 (2009)

Ruprecht, M., Bionda, T., Sato, T., Sommer, M.S., Endo, T., Scheiff, S. On the impact of precursor unfolding during protein import into chloroplasts. *Molecular Plant*. 1-10 (2010)

Schleiff, E., Jelic, M., Soll, J. A GTP-driven motor moves proteins across the outer envelope of chloroplasts. *PNAS*. **100** 4604-4609 (2003a)

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

Schnell, D.J., Blobel, G., Keegstra, K., Kessler, F., Ko, K., Soll, J. A consensus nomenclature for the protein-import components of the chloroplast envelope. *Trends in Cell Biology*. **7** 303–304 (1997)

Schulz, A., Knoetzel, J., Scheller, H.V., Mant, A. Uptake of a fluorescent dye as a swift and simple indicator of organelle intactness: import-competent chloroplasts from soil-grown *Arabidopsis*. *Journal of Histochemistry & Cytochemistry*. **52** 701-704 (2004)

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

Smith, M.D. Protein import into chloroplasts: an ever-evolving story. *Canadian Journal of Botany*. **84** 531-542 (2006)

Smith, M.D., Fitzpatrick, L., Keegstra, K., Schnell, D.J. *In vitro* analysis of chloroplast protein import. In: *Current Protocols in Cell Biology*. Eds. Bonifacino, J.S., Dasso, M., Harford, J.B., Lippincott-Schwartz, J., Yamada, K.M. New York: John Wiley and Sons Inc. 11.16.1-11.16.21 (2002b)

Smith, M.D., Rounds, C.M., Wang, F., Chen, K., Afitlhile, M., Schnell, D.J. atToc159 is a selective transit peptide receptor for the import of nucleus-encoded chloroplast proteins. *Journal of Cell Biology*. **165** 323-334 (2004)

Smith, M.D. & Schnell, D.J. Chloroplast Protein Targeting: Multiple Pathways for a Complex Organelle. *Protein Movement Across Membranes*. Ed. J. Eichler. Landes Biosciences. 1-18 (2004)

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

Stengel, A., Benz, P., Balsera, M., Soll, J., Bolter, B. Tic62 redox-regulated translocon composition and dynamics. *Journal of Biological Chemistry*. **283** 6656-6667 (2008)

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

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

Thompson, W.W., and Whatley, J.M. Development of non-green plastids. *Annual Review Plant Physiology*. **31** 375-394 (1980)

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

Tranel,P.J., Keegstra, K. A novel, bipartite transit peptide targets OEP75 to the outer membrane of the chloroplastic envelope. *The Plant Cell*. **8** 2093-2104 (1996)

Tripp, J., Inoue, K., Keegstra, K., Froehlich, J.E. A novel serine/praline-rich domain in combination with a transmembrane domain is required for the insertion of atTic40 into the inner envelope membrane of chloroplasts. *The Plant Journal*. **52** 824-838 (2007)

Tsai, L.Y., Tu, S.L., Li, H.M. Insertion of atToc34 into the chloroplastic outer membrane is assisted by at least two proteinaceous components in the import system. *Journal Biological Chemistry*. **274** 18735-18740 (1999)

Tu, S-L., Chen, L-J., Smith, M.D., Su, Y-s., Schnell, D.J., Li, H-m. Import pathways of chloroplast interior proteins and the outer-membrane protein OEP14 converge at Toc75. *The Plant Cell*. **16** 2078-2088 (2004)

Uversky, V. What does it mean to be natively unfolded? *European Journal of Biochemistry*. **269** 2-12 (2002)

Viana, A.A.B., Li, M., Schnell, D.J. Determinants for the stop-transfer and post-import pathways for protein targeting to the chloroplast inner envelope membrane. *Journal of Biological Chemistry*. **285** 12948-12960 (2010)

Vojta, L., Soll, J., Bolter, B. Protein transport in chloroplasts – targeting to the intermembrane space. *FEBS Journal*. **274** 5043-5054 (2007a)

Vojta, L., Soll, J., Bolter, B. Requirements for a conservative protein translocation pathway in chloroplasts. *FEBS Letters***. 581** 2621-2624 (2007b)

Walker, D., Chaddock, A.M., Chaddock., J.A., Roberts, L.M., Lord, J.M., Robinson, C. Ricin A chain fused to chloroplast-targeting signal is unfolded on the chloroplast surface prior to import across the envelope membranes. *Journal of Biological Chemistry*. **271** 4082-4085 (1996)

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

Wang, F., Agne, B., Kessler, F., Schnell, D.J. The role of GTP binding and hydrolysis at the atToc159 pre-protein receptor during protein import into chloroplasts. *Journal of Cell Biology*. **183** 87-99 (2008)

Weibel, P., Hiltbrunner, A., Brand, L., Kessler, F. Dimerization of Toc-GTPases at the chloroplast protein import machinery. *Journal of Biological Chemistry*. **278** 37312-37329 (2003)

Whitmore, L. and Wallace, B.A. DICHROWEB: an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research* 32:W668-673 (2004)

Whitmore, L. and Wallace, B.A. Protein Secondary Structure Analyses from Circular Dichroism Spectroscopy: Methods and Reference Databases. *Biopolymers.* **89**, 392-400 (2008)

Wienk, H.L.J., Czisch, M., de Kruijff, B. The structural flexibility of the preferredoxin transit peptide. *FEBS Letters.* **453** 318-326 (1999)

Wise, R.R. The Diversity of Plastid Form and Function. In: The Structure and Function of Plastids. Eds. R.R. Wise, and J.K. Hoober. Dordrecht: Springer. 3-26 (2006)

Yan, J., Wang, J., Zhang, H. An ankyrin repeat-containing protein plays a role in both disease resistance and antioxidation metabolism. *The Plant Journal*. **29** 193-202 (2002)

Yan, X., Khan, S., Hase, T., Emes, M.J., Bowsher, C.G. Differential uptake of photosynthetic and non-photosynthetic proteins by pea root plastids. *FEBS Letters*. **580** 6509-6512 (2006)

7. Appendix

Appendix I – Colony Screening by PCR

pET21a:atToc132A transformed DH5α

PCR reactions were set up as per manufacturer's specifications (New England Biolabs) for ThermoPol *Taq* polymerase to a final volume of 15 µl. Primers used in the amplification of the atToc132A insert for subcloning, atToc132ANoHis.NdeI.S - (5'-ATAGTTATA**CATATG**GGAGATGGGACCGAG-3') and atToc132ANoHis.XhoI.AS - (5'-TTCTA**CTCGAG**T*CTA*AAGACCTGCTGGACG-3'), were used to amplify the same insert from positive, ampicillin resistant colonies. A portion of positive colonies were re-plated on LB-amp plates (50 µg/ml) and another portion added to PCR reactions to supply template DNA. The 35-cylce program was set up according to the manufacturer's specifications: 1 minute annealing phase at 58^0C, 1 minute 23 second (1 min/kb) extension phase at 72^oC. The colonies corresponding to reactions that showed amplification of the atToc132 insert were positive transformants.

Appendix II – Addition of single protein CD spectra

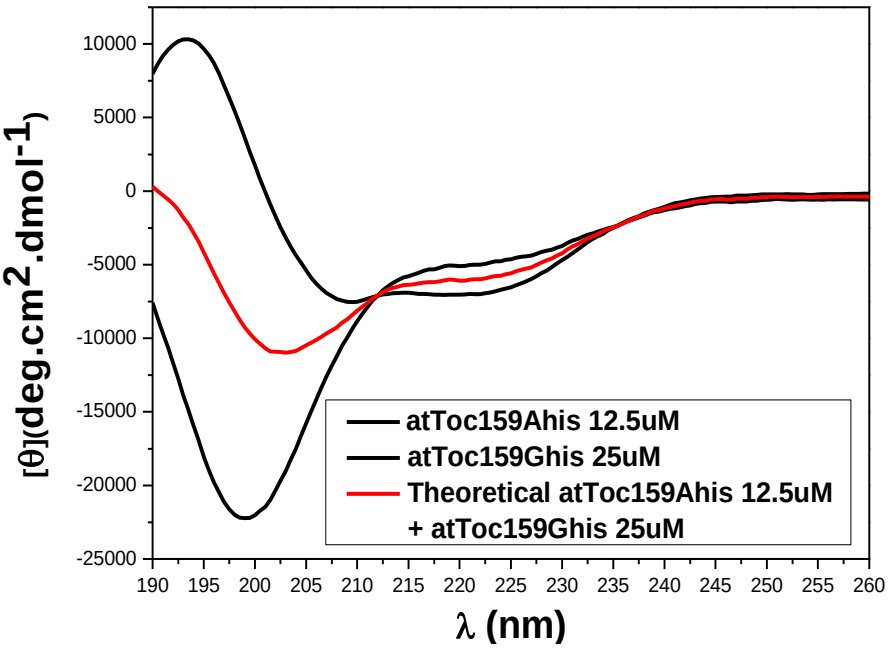


Figure 13. The addition of two single protein spectra (black traces) to produce a spectra that that is representative of the mixture of those two proteins (red trace) where no interaction or changes in secondary structure occur.

Appendix III – Concentration-Dependent Structural Change of atToc159G by Circular Dichroism

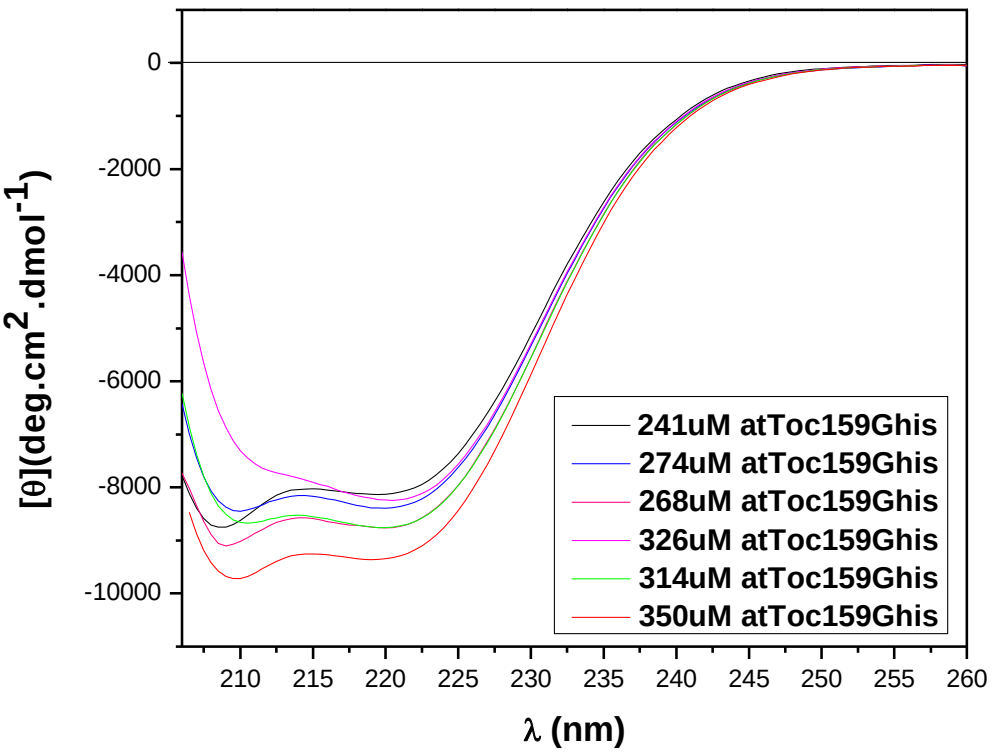


Figure 14. Concentration-dependence of secondary structure of atToc159G measured by CD. High concentration samples of atToc159G$_{His}$ were measured by CD in the Far-UV range. The change in concentration is not necessarily proportional to the change in spectra intensity suggesting that there may be some association between monomers occurring at high concentrations.