2013

TOC COMPLEX FORMATION: AN INVESTIGATION OF THE INTERACTIONS GOVERNING TOC COMPLEX COMPOSITION AND ASSEMBLY

Steven R. Siman
sima9760@mylaurier.ca

TOC COMPLEX FORMATION: AN INVESTIGATION OF THE INTERACTIONS GOVERNING TOC COMPLEX COMPOSITION AND ASSEMBLY

by

Steven Ronald Siman

BSc., Biology and Chemistry, Wilfrid Laurier University, 2009

THESIS

Submitted to the Department of Biology

Faculty of Science

in partial fulfillment of the requirements for the

Master of Science in Integrative Biology

Wilfrid Laurier University

2012

I hereby declare that I am the only author of this thesis.

I understand that my thesis may be made electronically available to the public.

Abstract

Chloroplast-destined preproteins are translated in the cytosol, and post-translationally targeted to and translocated across the double envelope membrane of the chloroplast by the coordinated activities of two translocon complexes: the Translocons at the Outer and Inner envelope membrane of the Chloroplast (TOC and TIC, respectively). In the model organism *Arabidopsis thaliana* the core TOC components include two families of GTPase receptors: TOC159 (atToc159, -132, and -120) and TOC34 (atToc33 and -34). These receptor families are hypothesized to assemble into distinct complexes and recognize transit peptides present on the N-terminus of chloroplast-destined preproteins. The GTPase domains of the TOC159 and TOC34 family members are hypothesized to interact in such a way that structurally and functionally distinct TOC complexes are formed. These distinct complexes are thought to have specificity for different subsets of preproteins.

Chloroplasts must differentiate between different subsets of proteins because they are needed in different amounts during various stages of chloroplast biogenesis. This investigation examines the propensity for atToc33 and atToc34 to associate with atToc159 or atToc132, how these interactions affect TOC complex formation, as well as what protein domains are conferring this preference. *In vitro* competitive chloroplast targeting assays, in which the GTPase domains of atToc33 or atToc34 are used as competitors for targeting of atToc159 or atToc132 to chloroplasts, and *in vitro* solid-phase binding assays, in which the GTPase domains of atToc33 or atToc34 are used as "bait" to test interactions with "prey" atToc159 or atToc132 are used to characterize these interactions. In order to study the influence of the highly divergent A-domain, these

associations are also being investigated using A-domain deletion mutants, atToc159GM and atToc132GM as well as A-domain swapped mutants 159A132GM and 132A159GM. This investigation has revealed that the mechanisms governing TOC GTPase interactions in *Arabidopsis* may be dictated by the A-domain of atToc132 and the G-domain of atToc159, thereby giving insight into how key TOC components are assembled into distinct TOC complexes at the chloroplast surface. Distinct complexes are responsible for the critical identification and import of different subsets of preproteins, all of which are necessary for plant growth and development.

Acknowledgements

I would like to thank Dr. Matthew Smith for getting me interested in research, and for his guidance and advice throughout my time as a graduate student. Thank you to my other thesis committee members Dr. Allison McDonald and Dr. Barbara Moffatt for their help and advice throughout the different stages of this project. Thank you to Siddhartha Dutta, Kyle Weston and Tuan (Patrick) Hoang as well as all other members of the Smith lab who were very generous with their time when I needed guidance and when learning new techniques. Thank you to Yi Chen and Jaime Ottaway whose work forms the building blocks of this project, and is used herein. Thank you to Lynn Richardson and Dr. Danny Schnell for providing some of the constructs used for making radiolabelled proteins, and to Kyle Weston and Yi Chen for producing the G-domains of atToc33 and atToc34, respectively. Thank you to Gena Braun and Jiangxiao Sun for their training and for helping whenever there were problems with lab equipment. Thank you to Dr. Tristan Long for his help with statistical analysis, and Christian Huynh for his help in designing successful primers. Six years and two degrees later, I would like to thank the Biology Department of Wilfrid Laurier University for a great ride. I have had wonderful experiences, met life-long friends, and this chapter of my life will not soon be forgotten.

A special thank you to my family and friends for all their love, support, and patience in getting me to where I am.

Dedicated to my dad, Ron Siman.

Table of Contents

List of Figures and Tables

List of Abbreviations

A-domain - Acidic domain of an *Arabidopsis* TOC GTPase protein

G-domain - GTPase domain of an *Arabidopsis* TOC GTPase protein

M-domain - Membrane anchor domain of an *Arabidopsis* TOC GTPase protein

FL159 - Full-Length atToc159

FL132 - Full-Length atToc132

159GM - A-domain truncated atToc159 (just the G- and M-domains)

132GM - A-domain truncated atToc132 (just the G- and M-domains)

159A132GM - A-domain swapped mutant consisting of the A-domain of atToc159 and the G- and M-domains of atToc132

132A159GM - A-domain swapped mutant consisting of the A-domain of atToc132 and the G- and M-domains of atToc159

33G - The GTPase domain of atToc33

34G - The GTPase domain of atToc34

IVT - *In Vitro* Translated product (radiolabelled with ^{35}S-Methionine)

*ppi*1 - Transgenic *Arabidopsis* "knockout" plant line; deficient in plastid protein import protein atToc33

*ppi*2 - Transgenic *Arabidopsis* "knockout" plant line; deficient in plastid protein import protein atToc159

*ppi*3 - Transgenic *Arabidopsis* "knockout" plant line; deficient in plastid protein import protein atToc34

1. Introduction

1.1 Plastids

Plastids are a class of organelles which can differentiate into a variety of sub-types from a common precursor known as a proplastid (Wise and Hoober, 2006). They are capable of transforming into these different forms interchangeably in response to stress (Nelson *et al.,* 2005), and to satisfy various needs within tissues (Bauer *et al.,* 2001; Inoue *et al.,* 2010). It is accepted that plastids arose from a cyanobacteria-like organism being engulfed by a eukaryotic cell via an endosymbiotic event (Wise and Hoober, 2006). Although plastids still house some of their own genetic information on a small circular genome, over time most of it has been transferred to the genome in the nucleus of the plant cell via a mechanism that is poorly understood (Leister, 2003). The majority of plastid proteins are now encoded by nuclear genes and synthesized in the cytosol. These translation products are known as protein precursors (or preproteins), as they carry a targeting sequence, and are not yet properly folded functional proteins (Agne and Kessler, 2009). Because the plastid is a membrane-bound organelle, the translocation of many different precursor proteins from the cytosol to the stroma of the plastid must be facilitated across a double membrane. Once translocated, protein precursors can either remain in the stroma or be targeted to one of the sub-compartments within the plastid (Aronsson and Jarvis, 2009).

Exposure to light, as well as developmental signals inherently present in proplastids, induce genetic signalling pathways that lead to photomorphogenesis of chloroplasts in green tissues (Bauer *et al.,* 2000). Chloroplasts are one subclass of plastid that are abundant as compared to other plastid types in the green tissues of plants. Their

most distinguished role is to harvest light energy from the sun via the process of photosynthesis, although they perform other critical non-photosynthetic functions as well, such as lipid and amino acid biosynthesis (Nelson *et al.,* 2005). Chloroplasts are used as a model for plastid protein import, as they have been well characterized, and are abundant and readily obtainable. A diagrammatic representation of a chloroplast illustrating its compartments is shown in Figure 1.

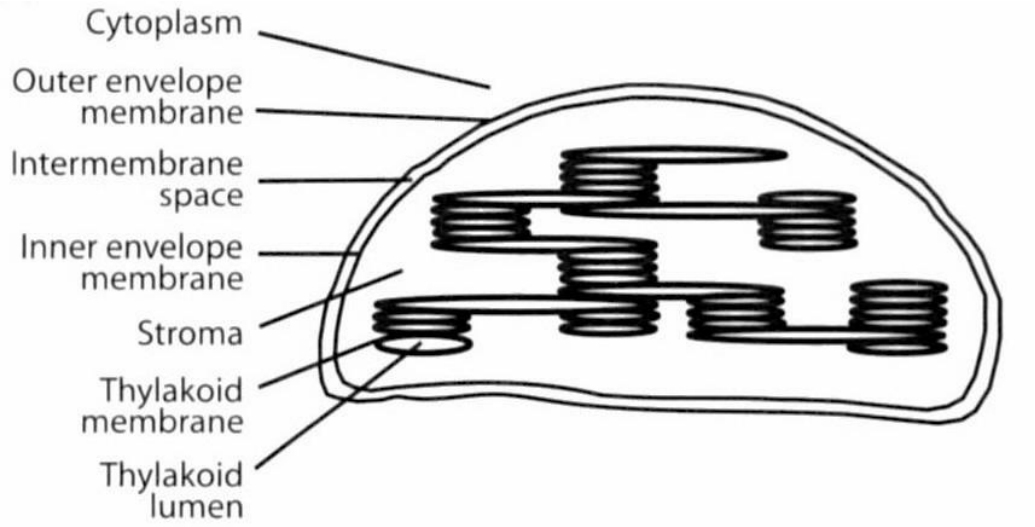


Figure 1 - Chloroplast General Structure. A schematic representation of the compartments of the chloroplast. (Smith, 2006)

1.2 Protein Import Into Chloroplasts

The envelope of a chloroplast is composed of two membranes (Figure 1); each of these is a bilayer consisting primarily of galactolipids, with a smaller proportion of phospholipids, and sulfolipids (Wise and Hoober, 2006) that effectively act as a barrier to

proteins, ions and other cellular constituents (Inoue, 2011). Plastid proteins that are encoded by the nuclear genome of the plant cell must be translated in the cytosol, post-translationally targeted to the plastid and translocated across the double membrane envelope (Smith, 2006). Embedded in the membranes are two translocon complexes whose activities are coordinated: the Translocon at the Outer membrane of the Chloroplast (TOC) and the Translocon at the Inner membrane of the Chloroplast (TIC) (Schnell *et al.,* 1997; Agne and Kessler, 2009).

Proteins that are encoded in the nuclear genome, but are destined for the chloroplast, contain additional amino acids which contain targeting information (Agne and Kessler, 2009). The targeting sequence, known as a transit peptide, is on the N-terminal end of what is called the precursor protein and allows the TOC and TIC complexes of the chloroplast to identify precursors and subsequently transport them across the membrane (Agne and Kessler, 2009). Transit peptide sequences are highly variable and can be anywhere from 13-146 amino acids in length (Zhang and Glaser, 2002). The sequence is cleaved upon import by a stromal processing peptidase (Richter *et al.*, 2005). The protein is then ready to be folded and become functional, or subsequently can be targeted to a chloroplast sub-compartment.

1.3 Role of the TOC Complex

The TOC complex is responsible for recognizing the transit peptides of precursor proteins destined for the chloroplast, leading to protein import. It works in conjunction with the TIC complex to move preproteins across the double membrane. While the core TOC proteins have been elucidated (Figure 2), there are still many unanswered questions

pertaining to the specific structure and mechanisms of activity of the complex (Bauer *et al.,* 2000; Smith *et al.,* 2002b). The research project described herein focuses on the TOC complex.

The current project focuses on the chloroplast translocon machinery of *Arabidopsis thaliana;* this plant has a short life cycle and relatively small genome (Meinke *et al.,* 1998) which has been sequenced and is publicly available for use (*Arabidopsis* Genome Initiative, 2000). It has become a model species for much research in the field of plant biology, including molecular studies of mechanisms within the plant (Salinas and Sanchez-Serrano, 2006). Its short life cycle allows tissues to be grown for easy use in molecular and biochemical techniques in ~14-21 days (on sterile agar plates, in optimal conditions). The efficient transformation that is possible with *Arabidopsis* (Clough and Bent, 1998) has provided an opportunity for many mutant lines to be established over time by various groups (Bauer *et al.,* 2001). This has given rise to many tools and approaches for studying cellular processes at the molecular level, which often involves introducing mutations, leading to altered versions of specific proteins, or the elimination of specific proteins altogether. For example, by creating "knockout" mutants, the role of a given protein or complex can be investigated by examining the consequences for the organism in its absence (Azpiroz-Leehan and Feldmann, 1997). Alternatively, mutated or altered versions of a given *Arabidopsis* protein can be introduced into a system either *in vitro* or *in vivo*.

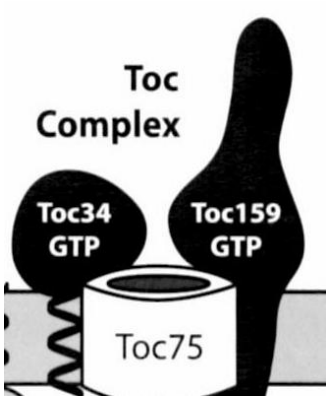


Figure 2 - Core TOC Components. A general representation of the core components of the TOC complex. The complex is known to be embedded in the outer envelope membrane of the chloroplast. TOC159 has four homologues in *Arabidopsis*: atToc159, atToc132, atToc120, and atToc90; TOC 34 has two *Arabidopsis* homologues: atToc33 and atToc34. (Smith, 2006)

1.4 Core TOC Components

atToc75 is the channel protein that provides the conduit for preproteins to cross the outer membrane and subsequently be transferred to the TIC complex (Agne and Kessler, 2009). It has been shown to be present in all forms of the TOC complex (Keegstra and Cline, 1999; Ivanova *et. al.*, 2004).

The atTOC159 family proteins (Figure 3a) are known to consist of three distinct domains: the C-terminal Membrane domain (M-domain) anchors the protein to the outer membrane of the chloroplast envelope by an unknown mechanism, the GTPase (G-domain) is thought to play a key role in identifying and initiating translocation of

precursor proteins (Schleiff *et al.*, 2003), and the N-terminal Acidic domain (A-domain) is intrinsically unstructured and is the largest of the three domains in all members of the atTOC159 family (Richardson *et al.*, 2009). As represented schematically in Figure 3a, the G- and M-domains within this protein family are the most similar (highly conserved), whereas the A-domains are the most divergent (Kubis, *et al.*, 2004). Targeting of atToc159 to the chloroplast outer envelope has been shown to be dependent on an interaction with the atToc34 family, and proper insertion also requires atToc75 to be present (Wallas *et al.*, 2003). The concerted preprotein insertion mechanism is GTP-dependent and this is discussed in further detail in section 1.6.

Members of the atToc34 family (Figure 3b) consist of only two distinct domains: the C-terminal M-domain and the GTPase-domain (Richardson *et al.*, 2009). They are tail-anchored proteins capable of self-dependent insertion (which has been shown to not be true of all tail-anchored proteins) into the plastid outer envelope (Dhanoa *et al.*, 2010).

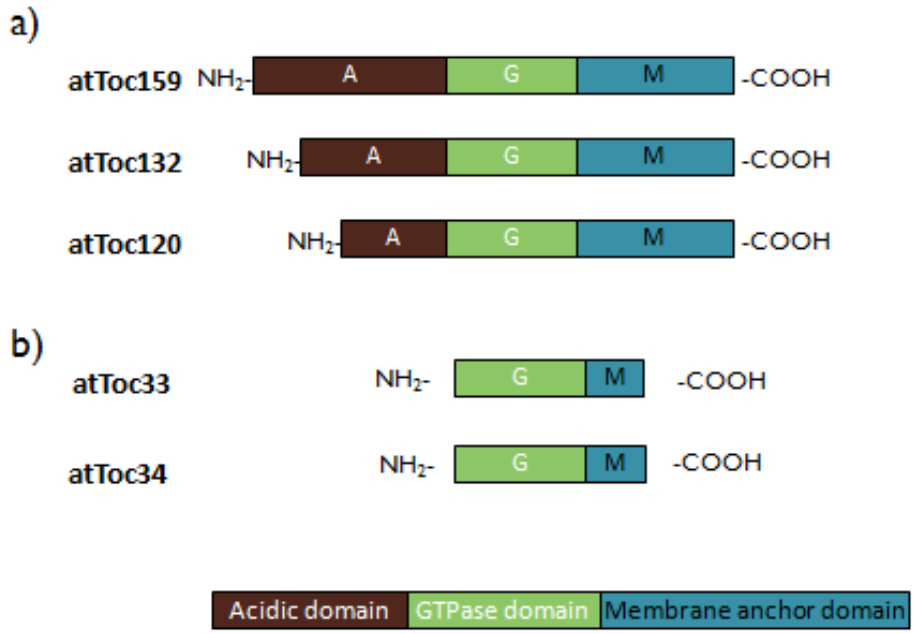


Figure 3 - Domains of *Arabidopsis* TOC GTPases. Domain structure of the TOC159 (a) and TOC34 (b) families of GTPase receptors in *Arabidopsis*. A, Acidic-domain; G, GTPase-domain; M, Membrane-domain.

1.5 TOC Complex Machinery

A given TOC complex (Figure 2) is comprised of members of several families of proteins which have been named to reflect their molecular weight (Schnell *et al.*, 1997). TOC159 is the largest protein in the complex, and is one member of a family of GTPase receptors (the TOC159 family) that are believed to be differentially present in structurally and functionally distinct TOC complexes (Ivanova *et al.*, 2004; Kubis *et al.*, 2004). In *Arabidopsis* the TOC159 family consists of: atToc159, atToc132, atToc120 and atToc90

(Bauer *et al.,* 2000; Ivanova *et al.,* 2004; Smith, 2006). This family of receptor proteins has been shown to be necessary for plastid biogenesis using an atToc159 knockout mutant and supplementing with the different TOC159 isoforms in *Arabidopsis* (Bauer *et al.,* 2000), and the different versions have been shown to be involved in the import of specific preproteins (Ivanova *et al.,* 2004; Kubis *et al.,* 2004; Lee *et al.,* 2009). It is important for the chloroplast to be able to recognize different types of proteins for different stages of biogenesis, and so that an appropriate response can be made to changing conditions. It has been suggested that the two major subsets of imported proteins to the chloroplast can be categorized into photosynthetic, and constitutive house-keeping proteins (Bauer *et al.,* 2000; Ivanova *et al.,* 2004; Smith, *et al.,* 2004); however, this model may be over-simplified (Agne and Kessler, 2009). The distinct targeting pathways (which are a product of distinct combinations of the GTPase receptor molecules) are hypothesized to allow for simultaneous import of proteins with different expression levels, and would minimize competition for import between unrelated subsets of proteins (Ivanova *et al.,* 2004).

While little is known about atToc90, it has been demonstrated by transgenic complementation studies that atToc132 is closely related to atToc120 (they also share 69% sequence identity, and uniform expression patterns relative to atToc159; Kubis *et al.,* 2004), and these versions of the GTPase protein form complexes distinct from those formed by atToc159 (Ivanova *et al.,* 2004). atToc132 and atToc120 have been observed together in a TOC complex, however atToc159 did not demonstrate interaction with either of atToc132 or atToc120, indicating that atToc159 exists in structurally distinct TOC complexes (Ivanova *et al.,* 2004). It has also been shown that atToc132/atToc120

and atToc159 associate preferentially (but not exclusively) with different versions of the atToc34 GTPase family of proteins (Ivanova *et al.*, 2004). In *Arabidopsis*, there are two proteins in the TOC34 family: atToc33 and atToc34. In a given TOC complex, atToc132/atToc120 show a preference for associating with atToc34, and atToc159 a preference for atToc33 (Figure 4;Ivanova *et al.*, 2004). These different forms of the TOC complex are believed to be both structurally and functionally distinct, having different affinities for various subsets of precursor proteins (Ivanova *et al.*, 2004; Kubis *et al.*, 2004). This specificity is presumably important for the plant because the plastid will require differing amounts of various types of proteins during the course of biogenesis, and respond to stresses, including changing environmental conditions. For example, when exposure to light stimulates photomorphogenesis a developing chloroplast must import a large number of photosynthetic proteins; continued import of other equally important but non-photosynthetic proteins is also important at this time, and therefore the chloroplast needs a mechanism to ensure that the massive influx of photosynthetic proteins does not overwhelm the ability to import other types of proteins. This is conceivably regulated by controlling the number and ratio of different TOC complexes that are present in the outer membrane at a given time.

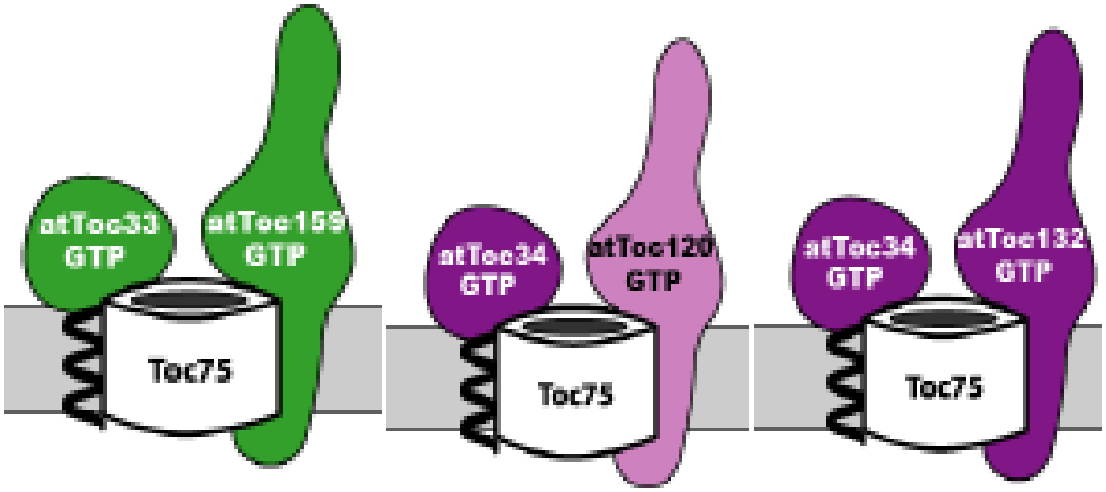


Figure 4 - TOC GTPase Receptor Preferred Associations in *Arabidopsis*.

1.6 TOC GTPase Domains (G-Domains)

Most GTPases are used by the cell to regulate a variety of signalling, synthesis and transport processes (Koenig *et al.*, 2008; Aronsson and Jarvis, 2011). A characteristic feature of GTPases is that their structure contains five loops (G1-G5) which function to bind and hydrolyze GTP (Koenig *et al.*, 2008). All five of these loops have been identified in the TOC GTPases (i.e. the TOC159 and TOC34 families; Sun *et al.*, 2002). A form of regulation employed by some GTPases is dimerization (though there are structural and functional differences among GTPase dimers from different systems; Koenig *et al.*, 2008). Dimerization has also been shown to play a role in TOC complex formation, although the exact function of dimerization has yet to be determined (Aronsson and Jarvis, 2011). TOC GTPases belong to the superclass of P-loop NTPases and the paraseptin subfamily of the TRAFAC family (Agne and Kessler, 2009).

The G-domains of the TOC GTPases are highly conserved (between 44-93% sequence identity; Jarvis, *et al.*, 1998; Bauer *et al.,* 2002; Kubis, *et al.,* 2004), and are of great importance in regulating and carrying out protein import (Aronsson and Jarvis, 2009). The GTPase domains of both the TOC159 and TOC34 families of receptor proteins have been shown to dimerize (with themselves, and with other TOC GTPases), and it is hypothesized that these interactions are critical to TOC complex assembly and therefore protein import (Kessler and Schnell, 2002).

1.6.1 The Role of TOC33/34 G-Domains in Assembly of Distinct TOC Complexes

It is believed that the interaction between the TOC34 family of proteins is regulated by GTP binding and hydrolysis (Agne and Kessler, 2009), as well as by precursor binding (Oreb *et al.*, 2011). atToc33 and atToc34 have been shown to have functional redundancies based on the observation that the two similar proteins atToc33 and atToc34 are both able to (at least partially) rescue *ppi*1 mutants (atToc33 deficient) (Jarvis *et al.*, 1998) and *ppi*3 mutants (atToc34 deficient) (Constan *et al.*, 2004). Changes in phenotype of the rescued mutant lines, however, suggest functional differences between atToc33 and atToc34 leading to the proposed existence of at least two distinct forms of TOC complex (Constan *et al.,* 2004). It has been demonstrated that the two isoforms atToc33 and atToc34 can be found in distinct TOC complexes, and that their preference for dimerizing with atToc159 or atToc132/120, respectively, is indicative of these proteins contributing (via G-domain interactions) to the differentiation of distinct targeting pathways to the chloroplast (Ivanova *et al.*, 2004). The amount of each different form of structurally and functionally distinct TOC complex present at the chloroplast

surface at a given time would presumably be dictated in part by these interactions. It is also becoming apparent that in *Arabidopsis,* the TOC34 family members (atToc33 and atToc34) may have additional differences in terms of how their expression and activity are regulated (Gutensohn *et al.,* 2000; Jelic *et al.,* 2003).

One model suggests atToc33 homodimers (in the GDP loaded state) are disrupted by precursor binding, allowing one of the monomers to be activated by binding GTP (Oreb *et al.,* 2011). It has been suggested that this step may be required for one of the atToc33 monomers to be capable of interaction with atToc159 (Sommer and Schlieff, 2009) (Figure 5a). Indeed, it has been suggested that GTP could specifically disrupt an atToc34-precursor interaction, perhaps encouraging association with atToc132/120 (Gutensohn *et al.,* 2000). If this phenomenon occurs in the atToc132/120-atToc34 system and not the atToc159-atToc33 system, it could suggest one of many potential mechanisms of regulating TOC complex formation, (Figure 5b), as atToc34-atToc132/120 may be more sensitive to GTP levels due to the receptor-precursor interaction being destabilized, presumably by a conformational change) than would atToc33-atToc159. What happens to the second monomer in this model that does not take the precursor protein is still unknown (Oreb *et al.,* 2011).

It has been suggested that, in general, the TOC159 family of proteins is unable to interact with dimers of the TOC34 family (Sommer and Schleiff, 2009) based on experiments that showed that TOC34 family interactions with precursor proteins can only be observed in the presence of GDP, while GTP promotes precursor transfer to a TOC159 family protein (Becker *et al.,* 2004, Kouranov and Schnell, 1997, Oreb *et al.,* 2011, Sommer and Schleiff, 2009). Taken together, this supports the aforementioned

model.

It is also known that heterodimerization between the TOC34 and TOC159 families occurs via the G-domains (Bauer *et al.*, 2002), however, it is accepted that heterodimerization between members of the TOC34 family (atToc33, -34) also occurs (Kessler and Schnell, 2002), and is not addressed by the model described above. It could be that this is perhaps one of various regulatory systems influencing TOC complex assembly and by extension protein import.

Jelic *et al.* (2003) demonstrated that atToc33 can be phosphorylated, while atToc34 cannot. These authors showed that the phosphorylation of atToc33 inhibits GTP binding, and as a consequence does not allow interaction with the precursor protein. Gutensohn *et al.* (2000) showed that atToc33 and atToc34 are found differentially throughout tissue types. Both Jelic *et al.* (2003) and Gutensohn *et al.* (2000) suggest that atToc33 and -34 exhibit some preference for precursor binding, suggesting each is important for import of different subsets of proteins.

By conducting experiments with all combinations of the G-domains of atToc33 and atToc34 with all members of the TOC159 family in *Arabidopsis*, the current project aims to shed more light on the intricate mechanisms governing TOC complex assembly and what role is played by both of atToc33 and -34 in this process.

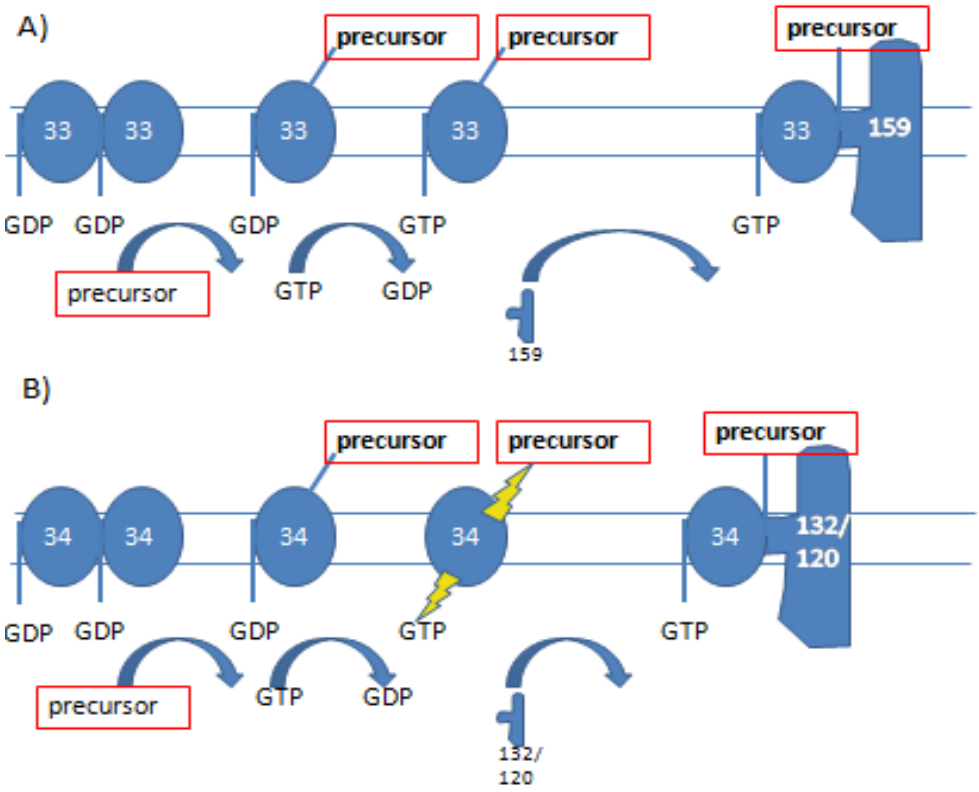


Figure 5 - Hypothesized Modes of Action in Preprotein Import for the Members of the TOC33 Family in *Arabidopsis*. A) atToc33 dimer is disrupted by binding with a compatible protein precursor, allowing one monomer to exchange GDP for GTP, which subsequently allows for potential interaction with atToc159. B) atToc34 dimer is disrupted by binding with a compatible protein precursor, allowing one monomer to exchange GDP for GTP, which has been suggested (Gutensohn *et al.*, 2000) to disrupt atToc34-precursor binding, possibly promoting interaction with atToc132/120. According to this model, it may be easier to get atToc132/120 to interact with atToc34 than atToc159 to interact with atToc33 due to the atToc34-precursor interaction being destabilized. If this is the case, a "GTP switch" (subtle change in GTP concentration; Bauer *et al.*, 2002) could have different effects on the two systems (specifically the 132/120-34 system might interact at a lower concentration of GTP than would the 159-33 system).

1.6.2 The Role of G-Domain Interactions Between TOC159 and TOC34 Families

Elucidating how the differences between atToc33 and -34 influence regulation of TOC complex assembly and preprotein import is made challenging by also considering interaction with the different isoforms atToc159, -132, and -120. Although atToc33 is hypothesized to have a preference for associating with atToc159, and atToc34 a preference for atToc132 and -120, such interactions are not believed to be exclusive (Ivanova *et al.,* 2004). While interactions between the two TOC GTPase families are perhaps the most studied, all combinations of interaction within the TOC159 and TOC34 families are believed to occur (Agne and Kessler, 2009). atToc33 and atToc34 can form homo- or heterodimers, and both of atToc132 and atToc120 have been demonstrated to interact with themselves and with one another, and can presumably be found together in a TOC complex (Ivanova *et al.,* 2004). However, atToc159 was not found in the same complex with either of atToc132 or atToc120 (Ivanova, *et. al.* 2004). Both families of GTPase receptors (TOC159, TOC34) have been shown to interact with precursor proteins (Kouranov and Schnell, 1997; Jarvis *et al.*, 1998; Ivanova *et al.*, 2004).

There is evidence that atToc159 is able to switch between a soluble and integral membrane form and that this switch is mediated by the G-domain of the protein (Bauer *et al.,* 2002; Hiltbrunner *et al.,* 2001). In addition, Lung and Chuong (2011) recently observed both Toc159 and Toc132 in cytosolic and membrane-associated forms in the *Bienertia sinuspersici* system. Although the existence of the soluble form is still contested by some (Soll and Schlieff, 2004), the insertion of atToc159 in the membrane occurs via interaction with atToc33/34 (Wallas *et al.*, 2003), which provides a potential mechanism for TOC assembly to support the model (Soll and Schlieff, 2004). It has been

shown that mutant atToc159 which is unable to bind GTP remains trapped in the cytosol, unable to become localized to the membrane (Bauer *et al.*, 2002). This leads to the inference that atToc159 depends on a GTP-regulated "switch" that allows it to associate with a TOC GTPase receptor at the membrane (Bauer *et al.*, 2002). There is also evidence that atToc33 reacts differently to an increase in GTP than does atToc34 *in vitro*, specifically that atToc33 binds precursor proteins more strongly as GTP increases, whereas atToc34 continues to have weak binding affinity for precursors as GTP increases (Gutensohn *et al.*, 2000). It has even been suggested that GTP could act to disrupt the atToc34-precursor interaction (Gutensohn *et al.*, 2000) (Figure 5). Conceivably then, the aforementioned GTP-regulated "switch" could have a different effect on the atToc159-atToc33 system than it does on the atToc132/120-atToc34 system.

1.7 Acidic Domain (A-Domain)

The function of the A-domain (only present in the TOC159 family) is not explicitly known (Smith *et al.*, 2002b); however, it is known to be intrinsically unstructured (Richardson *et al.*, 2009), extremely susceptible to protease degradation (Bölter *et al.*, 1998; Chen *et al.*, 2000), and is thought to play a role in determining the specificity of precursor binding (Inoue *et al.*, 2010; Dutta and Smith, unpublished data). While the G-domain is known to interact with preproteins (Bauer *et al.*, 2002), it is not believed to contain a preprotein sorting signal *per se* (in other words, it is not believed to confer specificity for preprotein interaction; Lung and Chuong, 2012). This is perhaps where the A-domain becomes of particular importance, although the exact sorting signal

(or mechanism by which the A-domain confers specificity to particular subsets of precursor proteins) is still a matter of some debate (Lung and Chuong, 2012).

Recent work by Lynn Richardson and Yi Chen, previous graduate students in Dr. Smith's lab, has suggested that the A domain plays a key role in the targeting of TOC159 isoforms to chloroplasts, and consequently TOC complex assembly. Specifically, it was shown that targeting efficiency of atToc132 to chloroplasts is increased when the A domain is removed, whereas deletion of the A-domain from atToc159 has no effect on targeting (Figure 6, from Richardson, 2008). This suggests that the atToc132 A-domain influences how much of the protein is able to bind at the chloroplast surface. To confirm the importance of the A domain in targeting atToc132 to chloroplasts, the atToc132 A-domain deletion mutant (132GM) was targeted to chloroplasts in the presence of increasing concentrations of the 132 A-domain added *in trans* (Figure 7, from Chen, 2011). In agreement with the data seen in Figure 6 (from Richardson, 2008), the A-domain inhibits targeting of atToc132. To investigate whether the A-domain imposes its effect on atToc132 targeting via atToc33 or atToc34, the experiment was repeated with the atToc33 & -34 knockout mutants, *ppi*1 and *ppi*3, respectively (Figure 8). When 132GM (A-domain deletion) is targeted to *ppi*3 chloroplasts (lacking atToc34), inhibition is observed (Figure 9, from Chen, 2011). In contrast, less inhibition is observed when the experiment is repeated with *ppi*1 chloroplasts (lacking atToc33) (Figure 9, from Chen, 2011). These data suggest that the 132 A-domain specifically inhibits atToc132 interaction with atToc33.

In order to gain more evidence to support the hypothesized role of the atToc132 A-domain, *in vitro* solid phase binding assays were used to confirm the relative strength

of interactions between full length atToc132 and 132GM with atToc33/34 (Chen, 2011; Ottaway, 2012). The A-domain deletion mutant (132GM) bound with much higher efficiency to atToc33G than did full-length atToc132 (Figure 10, from Ottaway, 2012), presumably because the A-domain inhibited the interaction. However, removal of the A-domain has less of an effect on the strength of interaction between atToc132 and atToc34G (Figure 11, from Ottaway, 2012). These data support the theory that the atToc132 A-domain plays a role in preventing interaction with atToc33. The current project not only repeats these experiments, but extends it to also test interactions between atToc33/34 and the A-domain swapped mutant proteins 159A132GM and 132A159GM.

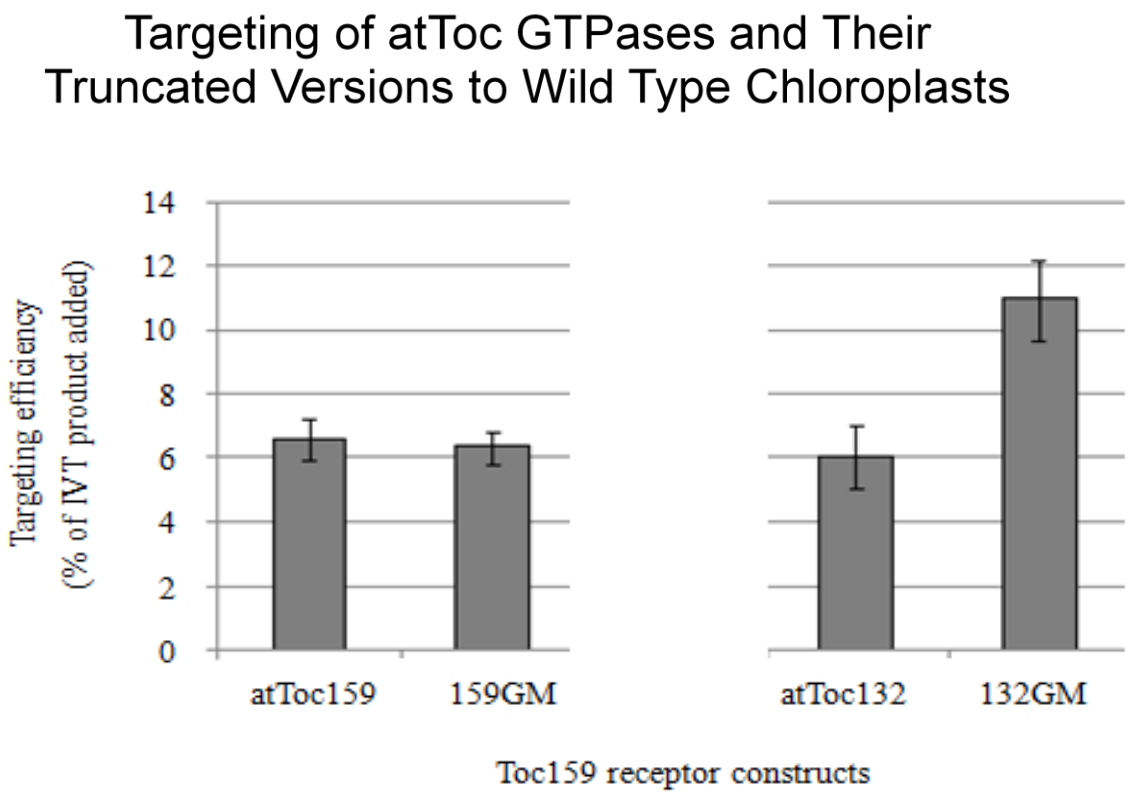


Figure 6 - Total targeting efficiency of full-length and A-domain truncation mutants (GM domains) of atToc159 and atToc132 to wild type chloroplasts using *in vitro* chloroplast targeting assays. (Richardson, 2008)

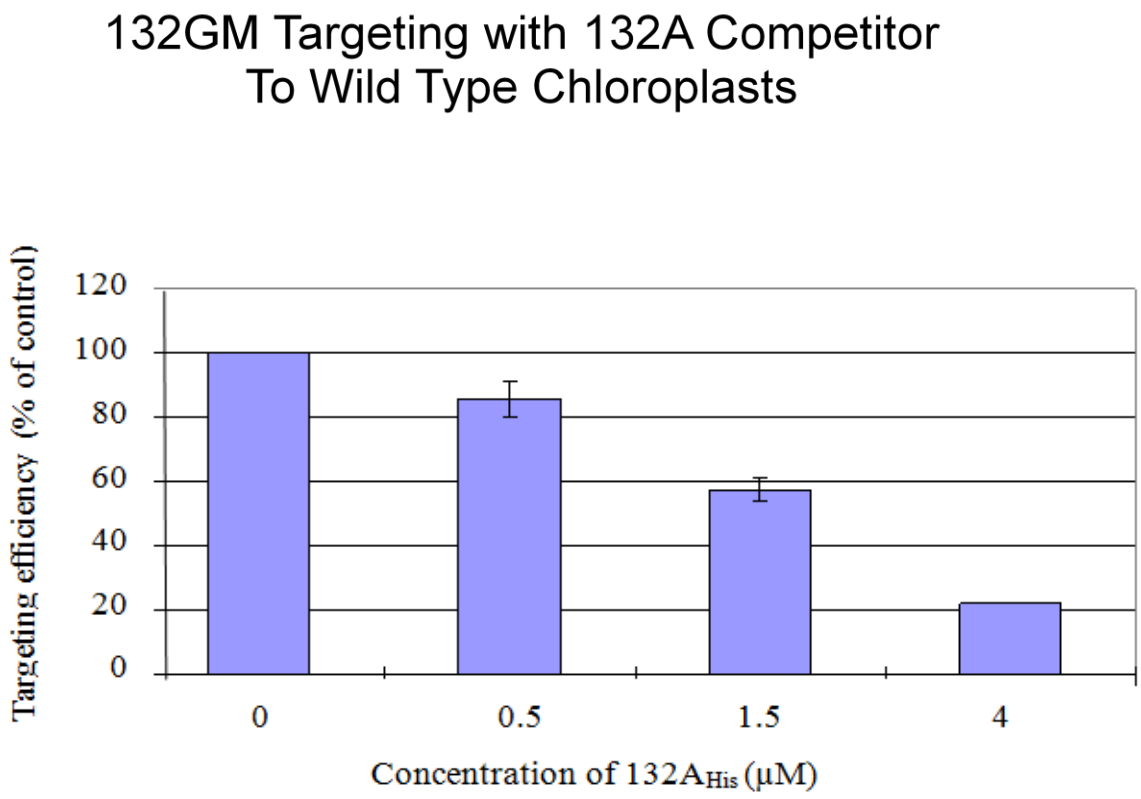


Figure 7 - Targeting of truncated versions (GM domains) of atToc132 to wild type chloroplasts in the presence of increasing amounts of atToc132 A domain. (Chen, 2011)

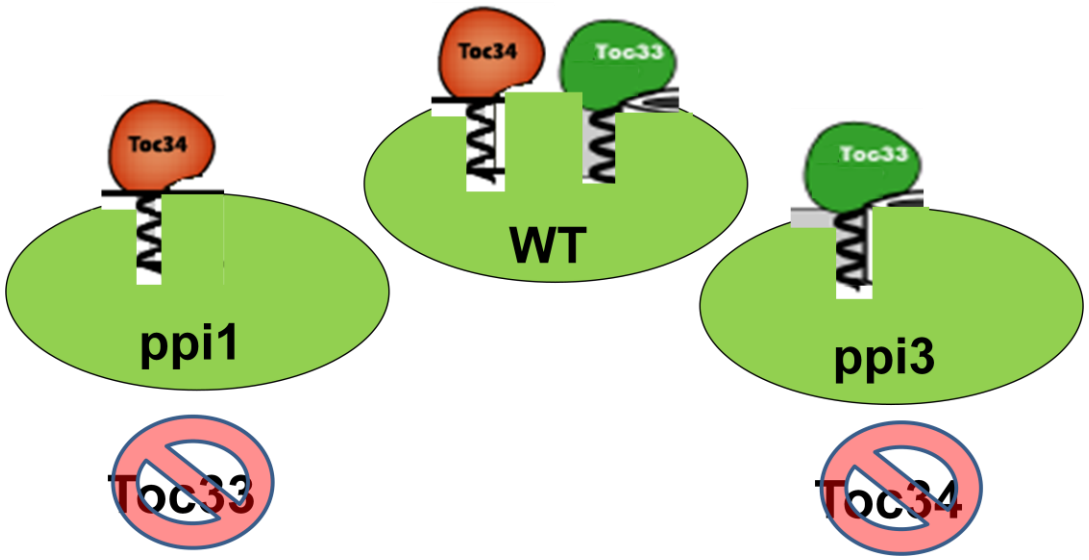


Figure 8 - Diagrammatic Representations of *ppi*1 and *ppi*3 Deletion Mutants.

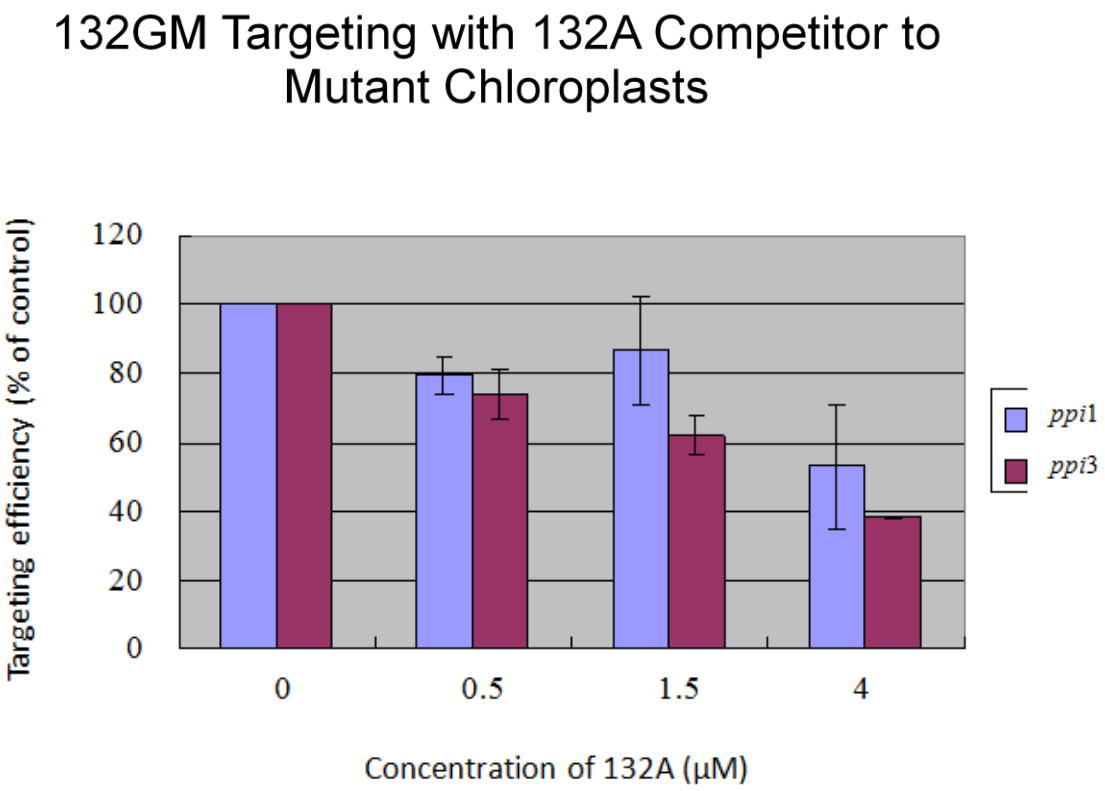


Figure 9 - Targeting of truncated versions (GM domains) of atToc132 to *ppi*1 and *ppi*3 mutant chloroplasts in the presence of increasing amounts of atToc132 A domain. (Chen, 2011)

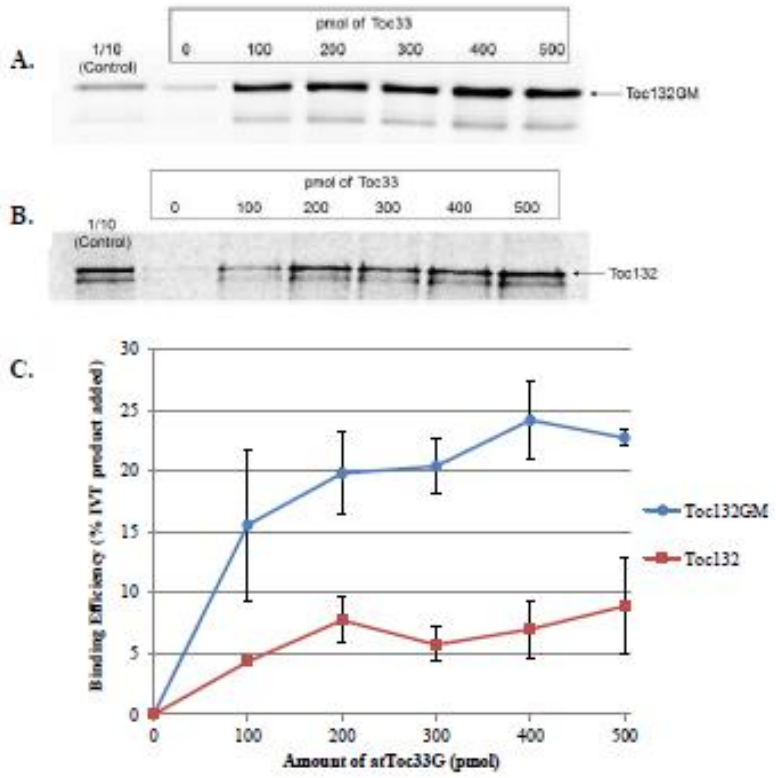


Figure 10 - Comparison of atToc132 and 132GM Interaction with 33G. *In vitro* solid phase binding assay, using atToc33G as bait, and atToc132GM (A) and atToc132 (B) as prey. C, quantitative analysis of the data presented in A and B. (Ottaway, 2012)

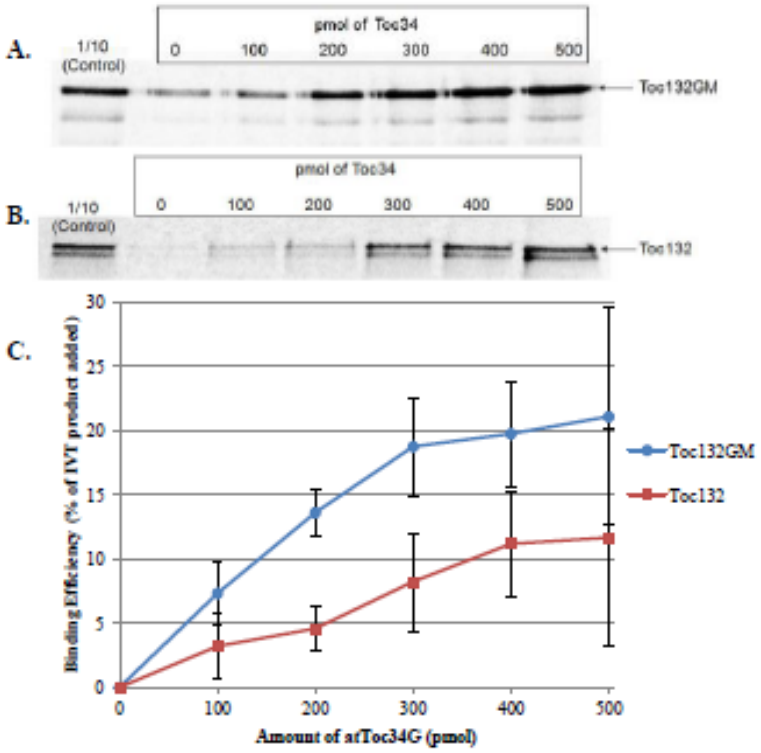


Figure 11 - Comparison of atToc132 and 132GM Interaction with 34G. *In vitro* solid phase binding assay, using atToc34G as bait, and atToc132GM (A) and atToc132 (B) as prey. C, quantitative analysis of the data presented in A and B. (Ottaway, 2012)

1.8 Research Objectives

By using both *in vitro* chloroplast targeting assays and *in vitro* solid-phase binding assays, the research described herein tests and compares the interactions between the *Arabidopsis* isoforms of the TOC159 family (atToc159, -132) and TOC34 family (atToc33, -34) in an effort to gain greater understanding of how TOC complex assembly occurs. Assays are also performed with A-domain deletion mutants (159GM, 132GM) and A-domain swapped mutants (132A159GM, 159A132GM).

Radiolabelling the native and mutated TOC159 isoforms, and comparing relative interactions with the GTPase domains of both of atToc33, -34, is intended to shed more light onto the mechanisms governing TOC complex structure and assembly. Testing these interactions both in the context of binding with the chloroplast in targeting assays (where endogenous native TOC family members are present at the chloroplast surface), as well as when removed from the the context of the chloroplast in solid-phase binding assays (only known amounts of proteins are present) is designed to show interaction preferences from different perspectives.

Testing interactions with both A-domain deletion and A-domain swapped mutants is intended to identify what role the A-domain plays versus the GTPase domain (and equally importantly, if this role is the same for the different isoforms of each GTPase family).

Objective: To examine the interactions among TOC GTPases in order to better understand TOC complex assembly.

Hypothesis: Specific interactions between members of the Toc159 and Toc34 GTPase families contribute to the formation of structurally and functionally distinct TOC complexes.

Specific aims: Use (i) *in vitro* chloroplast targeting assays and (ii) *in vitro* solid-phase binding assays to study the specificity of interactions between TOC GTPases, and which domains are responsible for conferring interaction preferences.

2. Materials and Methods

2.1 Plant Growth Conditions

Wild type *Arabidopsis thaliana* (ecotype Columbia) seeds were sterilized with 95% ethanol on a rotator for 5 minutes at room temperature, followed by treatment with 30% bleach containing 0.02% (v/v) Triton-X 100 for 20 minutes. Seeds were then washed 5-7 times with autoclaved Milli-Q water in a sterile flow hood, until the bleach was washed away completely. 30 mg/plate of seeds were sown onto 150 mm x 15 mm sterile plates on 0.8% (w/v) phytoblend media (Cassion, Cat.# PTP01) containing 1% (w/v) sucrose and 0.4% (w/v) Murashige and Skoog salt and vitamin mix (Caisson, Cat.#MSP01), pH 5.7. Plates were placed at 4°C for ≥48 hours in the dark to break the dormancy of the seeds. They were then placed in a growth chamber (Enconair, Bigfoot Series) and grown at 22°C under a 16:8 hour extended light cycle for 17-25 days. Light intensity in the chamber was measured at between 81-100 μM of photons /m^2/s.

2.2 *In vitro* Chloroplast Targeting Assays

2.2.1 Isolation of Intact Chloroplasts

Chloroplasts were isolated from 15-17-day-old plate-grown *Arabidopsis thaliana* seedlings as previously described (Brock *et al.,* 1993; Schulz *et al.,* 2004). Centrifuge rotors, tubes and all buffers were kept at 4°C. Approximately 150-200 g of tissue was separated from phytoblend media with a razor blade and homogenized using a PowerGen Homogenizer (Fisher Scientific) at setting 5 for ~15-20 seconds in pre-chilled grinding buffer (50mM Hepes-KOH, pH 7.5, 2 mM EDTA, 1 mM $MnCl_2$, 1 mM $MgCl_2$, 330 mM sorbitol, 1 mM ascorbic acid, 0.05% (v/v) Protease Inhibitor Cocktail (Sigma Cat.#

P9599). The homogenate was filtered through two layers of Miracloth (Calbiochem Cat.# 475855) into a chilled 500 ml centrifuge tube on ice and centrifuged at 1000x*g* for 8 min at 4°C (Beckman Coulter Avanti J-30I centrifuge; rotor JLA 10.5). The pellet, containing chloroplasts, was resuspended in ~8 ml of chilled grinding buffer by gently swirling while still holding on ice. The suspension was equally layered over two Percoll step gradients in chilled 50 ml centrifuge tubes consisting of a 7 ml lower 85% Percoll layer (85% (v/v) Percoll, 50 mM Hepes-KOH, pH 7.5, 330 mM sorbitol, 1 mM $MgCl_2$, 2 mM EDTA, 50 mM ascorbic acid, 0.2% (w/v) BSA) and an 8 ml upper 35% Percoll layer (35% (v/v) Percoll, 50 mM Hepes-KOH, pH 7.5, 2 mM EDTA, 1 mM $MnCl_2$, 1 mM $MgCl_2$, 330 mM sorbitol, 50 mM ascorbic acid). The gradients were centrifuged at 7700x*g* for 15 min at 4°C in a swinging bucket rotor (Beckman Coulter JS13.1) with slow acceleration and deceleration. The resultant upper band of broken chloroplasts (on top of the 35% Percoll layer) was aspirated. Intact chloroplasts at the interface between the Percoll layers were collected with a glass pipette, and diluted into HS buffer in another chilled 50 ml centrifuge tube (HS Buffer: 50 mM Hepes-KOH, pH 7.5, 330 mM sorbitol, 0.05% (v/v) Protease Inhibitor Cocktail (Sigma)). The intact chloroplasts were centrifuged at 1000x*g* for 6 min at 4°C. The resultant chloroplast pellet was resuspended in 300-600 µl chilled HS buffer by agitating gently over ice.

2.2.2 Quantification and Dilution of Intact Chloroplasts

The chlorophyll concentration of the isolated chloroplasts was measured at 652 nm (optimal optical absorption of chlorophyll *a* and *b*; Cary 50 Conc UV/Visible Spectrophotometer) by diluting chloroplast suspension 100x into 80% acetone as

described previously (Arnon, 1949). Chlorophyll content was determined as follows: chlorophyll concentration (mg/ml) = [A_{652}/36] x DF; where A_{652} is the absorbance measured at 652 nm and DF is the dilution factor of the intact chloroplast sample (DF=100 in this case). The freshly isolated intact chloroplasts were diluted to a final chlorophyll concentration of 1 mg/ml in HS buffer for use in import assays.

2.2.3 *In vitro* Transcription/Translation of Radiolabelled *Arabidopsis* TOC159 Homologues

Constructs used as templates (Table 1) for *in vitro* transcription/translation reactions were pET21d:atToc159 (Smith *et al.*, 2002b), pET21a:atToc132 (Bauer, *et al.*, 2000, Ivanova *et al.*, 2004), pET21d:159GM (Smith *et al.*, 2002b), pET21a:132GM (Smith *et al.*, 2002b), pET21d:159A132GM (Richardson, 2008), and pET21a:132A159GM (generous gift from D. Schnell, University of Massachusetts). Radiolabelled proteins were translated with the T_NT Coupled Reticulocyte Lysate System or T_NT Coupled Wheat Germ System (PromegaCat.#'s L4610 and L4140, respectively) in the presence of RNAsin (Promega# N251A). Briefly, 1 µg of plasmid DNA was used for a 50 µl reaction that included [^{35}S]Methionine (EasyTag EXPRES^{35}S^{35} Protein Lebeling Mix, Perkin Elmer, Cat .#NEG772002MC) in place of unlabelled methionine at 30°C for 1.5 hours, resulting in radiolabelled translation product. The reaction products were stored at -80°C until required.

Table 1 - Summary of Constructs Used for *in vitro* Transcription/Translation Reactions. The constructs, the proteins they code for and the terminology used herein to describe each are summarized.

Construct	**Radiolabelled Protein Product**	**Described as...**
pET21d:atToc159	atToc159	atToc159
pET21a:atToc132	atToc132	atToc132
pET21d:159GM	atToc159GM	159GM
pET21a:132GM	atToc132GM	132GM
pET21d:159A132GM	atToc159A132GM	159A132GM
pET21a:132A159GM	atToc132A159GM	132A159GM

2.2.4 *In vitro* Chloroplast Protein Targeting Assays

In vitro chloroplast protein targeting assays were performed as previously described (Smith *et al.*, 2002a). Each 100 µl targeting reaction contained chloroplasts equivalent to 50 µg of chlorophyll, 1 mM dithiothreitol (DTT), 5 mM ATP, 1 mM GTP, 10 mM methionine, 50 mM Hepes-KOH, pH 7.5, and 5 µl of radiolabelled protein (either atToc159, -132, 159GM, 132GM, 159A132GM, or 132A159GM; methionine is included in the reaction to out-compete residual excess ^{35}S-labelled methionine from IVT). A protein competitor [either atToc33G or atToc34G; atToc33G prepared by, and a generous gift from Kyle Weston (2011), atToc34G prepared by, and a generous gift from Yi Chen (2011)] was added in increasing concentrations to the reactions (0, 100, 200, 300, 400, 500 pmol) in order to test its effect on the targeting of the TOC159 orthologs to the chloroplast outer membrane. Prior to the addition of radiolabelled product, all other

components of the reaction mixture were equilibrated for 5 min in a waterbath at 26°C in the light. Targeting reactions were initiated with the addition of the radiolabelled product and allowed to continue for 30 min at 26°C in the light. The reaction was stopped by placing it on ice and adding 400 µl of ice cold HS buffer. Chloroplasts were re-isolated by layering the reaction mixture on top of 800 µl of 35% Percoll, and centrifuging for 5 min at 6000x*g*. Due to the fact that the chloroplasts do not form a firm pellet after one centrifugation, the upper portion of Percoll and excess buffer was aspirated down to ~100 µl. The bottom 100 µl contained a runny green pellet, and was transferred to 500 µl of HS buffer in a fresh tube. The chloroplasts were pelleted a final time by centrifuging again for 5 min at 6000x*g*. The supernatant was removed, and 25 µl of 2X SDS-PAGE sample buffer was added directly to the pellet and vortexed. The suspension was centrifuged for 30 sec at 10,000x*g* to pellet insoluble aggregates, and the supernatant was loaded onto an 8% SDS-PAGE gel (with a 4% stacking gel) for SDS-PAGE analysis. Control lanes were loaded containing the equivalent of 10% of the radiolabelled product that was added to each reaction. After the SDS-PAGE gels were run, they were stained (with Coomassie brilliant Blue) and destained, then dried for 1 hour using a Bio Rad 224 Gel Slab Dryer attached to a Thermo RVT400 refrigerated vapour trap and vacuum pump (Savant VP100). Radiolabelled proteins in the dried gels were detected by exposing to a phosphor screen (Kodak# SO230) in a casette for 5-7 days, and then imaged on Bio Rad Personal Molecular Imager FX phosphorimager. Quantity One 1-D Analysis software v4.6 (Bio Rad Laboratories Ltd.) was used to quantify the radio-labelled protein that associated with the chloroplast using the Volume Rect analysis tool.

A representative diagram of an *in vitro* chloroplast targeting assay is shown in Figure 12.

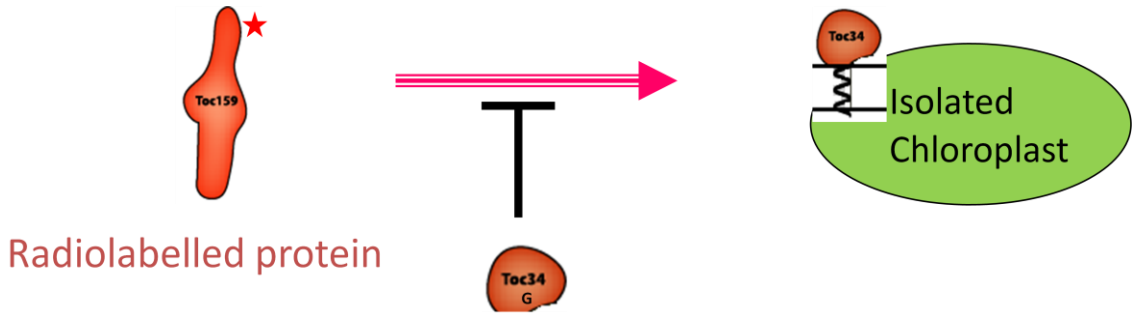


Figure 12 - *In vitro* Competitive Chloroplast Targeting Assay.

2.3 *In vitro* Solid-Phase Binding Assays

In vitro solid-phase binding assays were performed as previously described (Smith *et al.*, 2002b). Ni^{2+}-NTA resin (Novagen Cat. #71035) was washed 3X with Milli-Q water, and 3X with HMK Buffer [50 mM Hepes-KOH, pH 7.5, 5 mM $MgCl_2$, 1 mM THP (Tris-hydroxypropyl-phosphine; for reduction of sulfide bonds), 40 mM KOAc]. After each wash step, the resin was centrifuged at 10,000x*g* for 30 sec. The resin was resuspended in HMK buffer and aliquoted accordingly (equivalent of 20 µl of resin per reaction. Increasing concentrations (0, 100, 200, 300, 400, 500 pmol) of histidine-tagged atToc33G or atToc34G (Chen, 2011) were immobilized onto the resin in the presence of 1 mM GTP and incubated on a rotator for 30 min at room temperature. 198µl of HMKIT buffer (50 mM Hepes-KOH, pH 7.5, 5 mM $MgCl_2$, 1 mM THP, 40 mM KOAc, 10 mM imidazole, 0.1% Triton-X 100, 0.1 mM GTP) and 2µl of a particular radiolabelled protein (one of: atToc132, atToc132GM, atToc159A132GM, or atToc132A159GM) were added to each reaction, and placed back on the rotator for 30 min at room temperature. After

washing 3X with HMKIT buffer, proteins were eluted from the resin using 1x SDS-PAGE sample buffer containing 775mM imidazole (high imidazole concentration to compete for binding locations on the resin during elution). Samples were centrifuged at 10,000x*g* for 30 sec and run on an 8% (with 4% stacking) SDS-PAGE gel. The dried gels were put into a radiocassette with a phosphor screen (Kodak SO230) and exposed for 2-5 days, and imaged using a Bio Rad Personal Molecular Imager FX phosphorimager. Quantity One 1-D Analysis software v4.6 (Bio Rad Laboratories Ltd.) was used to quantify the amount of radio-labelled prey protein that was captured using the Volume Rect analysis tool.

A representative diagram of an in vitro solid-phase binding assay is shown in Figure 13.

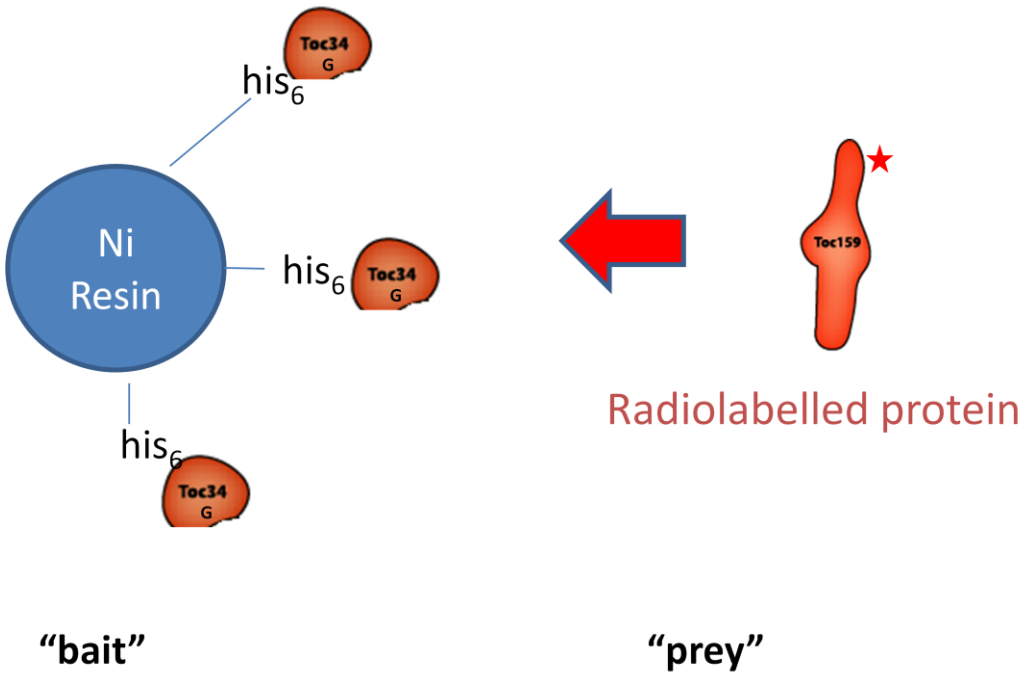


Figure 13 - *In vitro* Solid-phase Binding Assay.

3. Results

A current model states that atToc159 preferentially associates with atToc33 at the chloroplast surface, while atToc132 has a preference for atToc34, resulting in the formation of structurally and functionally distinct TOC complexes (Ivanova *et al.,* 2004; Kubis *et al.,* 2004). This model can be tested using *in vitro* competitive chloroplast targeting assays and *in vitro* solid-phase binding assays. According to the model, targeting of atToc159 to chloroplasts should be more strongly inhibited by the addition of excess G-domain of atToc33 (33G) than by the G-domain of atToc34 (34G). Conversely, it is expected that 34G will be a more effective competitor of atToc132 targeting than will 33G. The assay is based on previous studies that have demonstrated that an interaction between the GTPase domains of atToc159 and atToc33 is important for atToc159 targeting to chloroplasts (Smith *et al.*, 2002b; Bauer *et al.*, 2002; Kessler and Schnell, 2002), and an assumption that the inclusion of the GTPase domain of atToc33/34 added to the reaction in the soluble phase will interact with the atToc159/132 in the same way as the atToc33/34 already embedded in the membrane. In the case of *in vitro* targeting assays, binding by the competitor should preclude interaction with the corresponding receptor at the chloroplast, and prevent the targeting of atToc159/132 to the chloroplast surface which would occur through an interaction with the endogenous receptor proteins in the chloroplast envelope.

It should be emphasized that truncated mutants of atToc33 and atToc34 consisting only of the G-domains (with the C-terminal transmembrane domains deleted) were used as competitors to test interaction preferences. The transmembrane domain could not be included because its presence causes these proteins to be produced as inclusion bodies

rather than in the soluble phase in *E. coli*, which makes the proteins difficult to work with. In addition, presence of the transmembrane domain would complicate the competition assay, as the competitor would be targeted to the chloroplast and inserted into the membrane, which would affect assay results. Since the G-domain is the point of interaction for these proteins (Hiltbrunner *et al.,* 2001; Smith *et al.,* 2002b; Kessler and Schnell, 2002), it is necessary and sufficient to test interactions with the TOC159 family using these particular experimental approaches.

In vitro targeting and *in vitro* solid phase binding assays have been used in the current study to further examine the determinants of TOC complex assembly. The results of this project give some insight into the binding preferences among the TOC GTPases, and indicate that the interactions governing TOC complex assembly are likely more intricate than previously thought.

3.1 Targeting Assays

It has been demonstrated previously that atToc33G inhibits atToc159 targeting to isolated chloroplasts (Hiltbrunner *et al.,* 2001; Smith *et al.,* 2002b); however, atToc33 and atToc34 have not been compared, nor has the analysis been extended to include atToc132.

Recent evidence indicates that the A-domain of atToc132 may influence its relative affinity for atToc33 and atToc34 (Chen, 2011). Specifically, the evidence suggests that the A-domain inhibits the interaction of atToc132 with atToc33. This hypothesis is further tested in the current study.

In order to investigate the role of atToc33/34 in the targeting of atToc132/159 to

chloroplasts, this project used *in vitro* competitive chloroplast targeting assays, which involve monitoring the targeting of radiolabelled versions of the TOC159 family in *Arabidopsis* (in this case, atToc159 and atToc132) to isolated chloroplasts in the absence or presence of increasing concentrations of competitor proteins: the GTPase domains of atToc33 and atToc34. These experiments were designed to test the inherent binding preference exhibited by atToc132 for atToc34 and that of atToc159 for atToc33, and how these preferences are conferred. The role of the atToc132 A-domain in determining interaction specificity was of particular interest, as it may be a key to understanding assembly of structurally distinct TOC complexes.

Efficiency of targeting in the *in vitro* targeting assays was calculated from the intensity of bands in SDS-PAGE gels corresponding to radiolabelled proteins as measured using a phosphorimager. The intensity of a given band is representative of the amount of radiolabelled protein in the band. For each trial, the efficiency of targeting for the sample that contained zero competitor was set to 100%. Efficiency of targeting for samples in the presence of increasing concentrations (0.5, 2, and 4 μM) of competitor are represented as percentages relative to the control sample in the absence of competitor. As this is a relative comparison, error bars were calculated from the percentages and not from the raw data. In other words, an error bar represents the standard error by comparison of the relative data (represented by % difference from the zero value for a particular trial) for each trial. The reason standard error is calculated from the relative % differences for each trial and not from the absolute raw data is because (a) every chloroplast preparation is a little bit different, however, all potential protein interactions will be exposed to the same environment in a given trial and (b) the length of time a

phosphorimager screen is exposed to the dried gel affects the total intensity of the resultant bands.

Data sets for each particular data point were checked for normality and equal variance to confirm the data comprised a parametric system with the Shapiro-Wilks test and equal variance test, respectively, using "R" statistics software. Statistical relevance was determined by a standard t-test for parametric systems, and the Kolmogorov-Smirnov test for non-parametric systems. All systems were found to be statistically relevant, unless otherwise stated.

Each data set is based on values from between 4-6 trials (due to some data points having lost samples from a particular trial), and the error bars provided for each data point represent standard error, and account for how many trials were performed.

In targeting full-length atToc159 and atToc132 to chloroplasts in the presence of either 33G or 34G competitor protein, it was expected that 33G would be the more effective competitor of atToc159, and 34G the more effective competitor of atToc132, reflecting their previously observed preferences to associate with one another in distinct TOC complexes. Figure 14 shows that both atToc159 and atToc132 are strongly competed by 33G, while 34G is found to have no statistically relevant effect on the targeting of either protein.

In the case of 132GM, lacking the atToc132 A-domain, it was predicted that 33G and 34G would be roughly equivalent competitors of targeting to chloroplasts (in accordance with the hypothesis that it is the atToc132 A-domain which confers interaction preference of atToc132 for atToc34). In the case of 159GM, a similar result was predicted. As shown in Figure 15, 159GM targeting is strongly competed by 33G,

while any effect of competition from 34G was deemed statistically insignificant (indicating a strong preference of 159GM for 33G). Meanwhile, 34G was found to compete for the targeting of 132GM, whereas competition of this A-domain deletion mutant by 33G was deemed to be statistically insignificant (indicating a weak preference for 34G) (Figure 15). This result indicates that atToc159 does not require the presence of its A-domain to confer a preference for atToc33 (this is in agreement with Figure 6, from Richardson, 2008), and does not disagree with the hypothesis that the atToc132 A-domain hinders interaction with atToc33, as only a weak preference for 34G was observed.

The investigation was also extended to include testing interactions between 33G and 34G with A-domain swapped mutants 159A132GM and 132A159GM. It was expected (based on the hypothesis that the atToc132 A-domain inhibits interaction with atToc33) that targeting of 132A159GM would be inhibited to a greater extent by the presence of 34G competitor than by 33G. Contrary to the prediction, 34G was found to actually stimulate targeting to chloroplasts, while 33G did, in fact show competition (Figure 16a). While competition was observed for 159A132GM by 33G, once again (as with 132A159GM) 34G was found to stimulate targeting of this domain-swapped protein to the chloroplast (in fact, to an even greater extent than with 132A159GM), while 33G showed the expected competition (Figure 16b). It is reasonable to conclude that both the competition and stimulation of targeting that was observed were the result of interaction between the "competitor" (33G/34G) and the protein being targeted (domain-swapped proteins), which would suggest that the stimulation effect observed is a result of 34G shuttling atToc159 homologues to the chloroplast surface. Another possibility is that the

addition of 34G somehow stimulated the GTPase activity of the endogenous proteins, putting more of them in a GTP-bound state, which has previously been shown to be required for interaction with the TOC159 isoforms (Kouranov and Schnell, 1997; Becker *et al.,* 2004; Sommer and Schleiff, 2009; Oreb *et al.,* 2011) (See Figure 5).

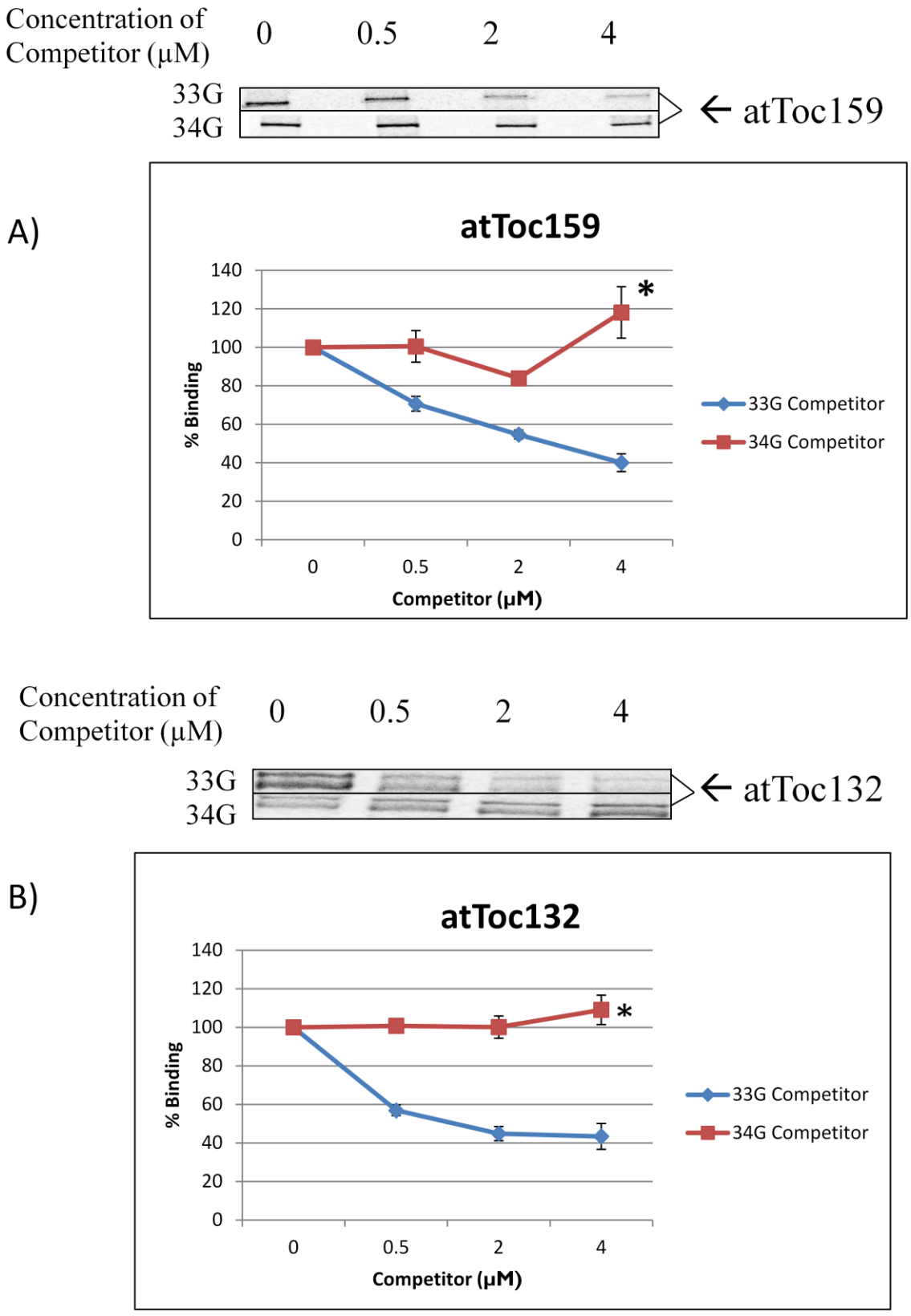


Figure 14 - *In vitro* Competitive Chloroplast Targeting Assays with Full-length Proteins

Figure 14 - *In vitro* Competitive Chloroplast Targeting Assays with Full-length Proteins. [^{35}S]-labelled atToc159 (A) and atToc132 (B) synthesized *in vitro* were targeted to isolated chloroplasts in the absence or presence of increasing concentrations of 33G or 34G as described in the Materials and Methods. Reactions were analyzed using SDS-PAGE, and targeting efficiency of atToc159 (A) and atToc132 (B) to wild type chloroplasts was analyzed by detecting radiolabelled proteins in dried gels using a phosphorimager. Efficiency of targeting is expressed as % Binding (percent of *in vitro* translated protein added to each reaction). Targeting efficiency is reflective of interaction with competitors. Gel images shown are representative of between 4-6 trials. Error bars represent standard error.

*indicates statistically insignificant effect.

(A) Data set was found to be normal using the Shapiro-Wilks test. Change was determined based on a standard t-test, having a p value >0.05 indicating no change.

(B) Data set was found to not be normal using the Shapiro-Wilks test, and could not be normalized via log transformation. Change was determined based on Kolmogorov-Smirnov test for non-parametric systems, having a p value <0.05 indicating no change.

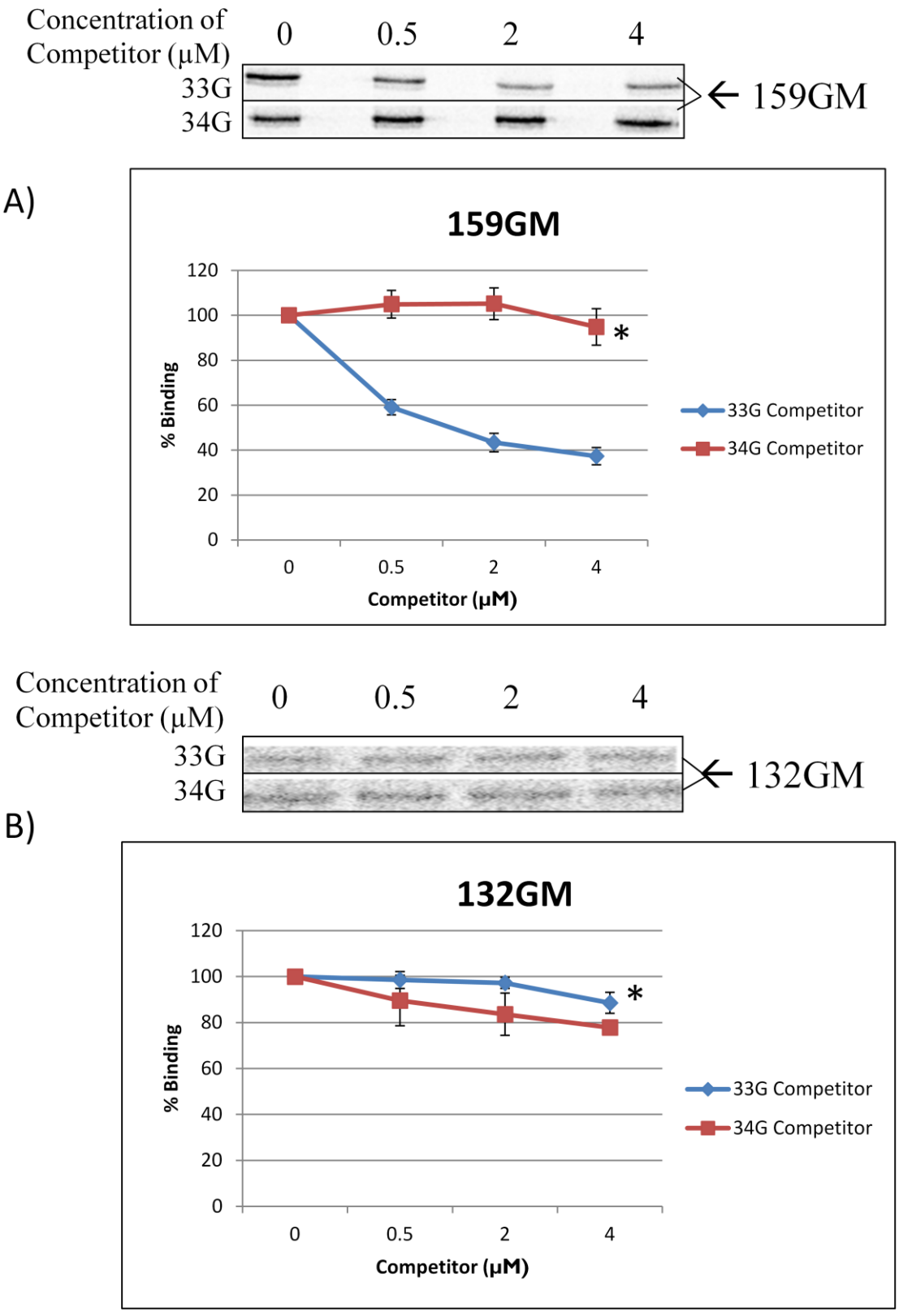


Figure 15 - *In vitro* Competitive Chloroplast Targeting Assays with A-Domain Truncated Proteins

Figure 15 - *In vitro* Competitive Chloroplast Targeting Assays with A-Domain Truncated Proteins. [^{35}S]-labelled 159GM (A) and 132GM (B) synthesized *in vitro* were targeted to isolated chloroplasts in the absence or presence of increasing concentrations of 33G or 34G as described in the Materials and Methods. Reactions were analyzed using SDS-PAGE, and targeting efficiency of 159GM (A) and 132GM (B) to wild type chloroplasts was analyzed by detecting radiolabelled proteins in dried gels using a phosphorimager. Efficiency of targeting is expressed as % Binding (percent of *in vitro* translated protein added to each reaction). Targeting efficiency is reflective of interaction with competitors. Gel images shown are representative of between 4-6 trials. Error bars represent standard error.

*indicates statistically insignificant effect.

(A) Data set was found not to be normal using the Shapiro-Wilks test, and was normalized via log transformation. Change was determined based on a standard t-test, having a p value >0.05 indicating no change.

(B) Data set was found not to be normal using the Shapiro-Wilks test, and was normalized via log transformation. Change was determined based on a standard t-test, having a p value >0.05 indicating no change.

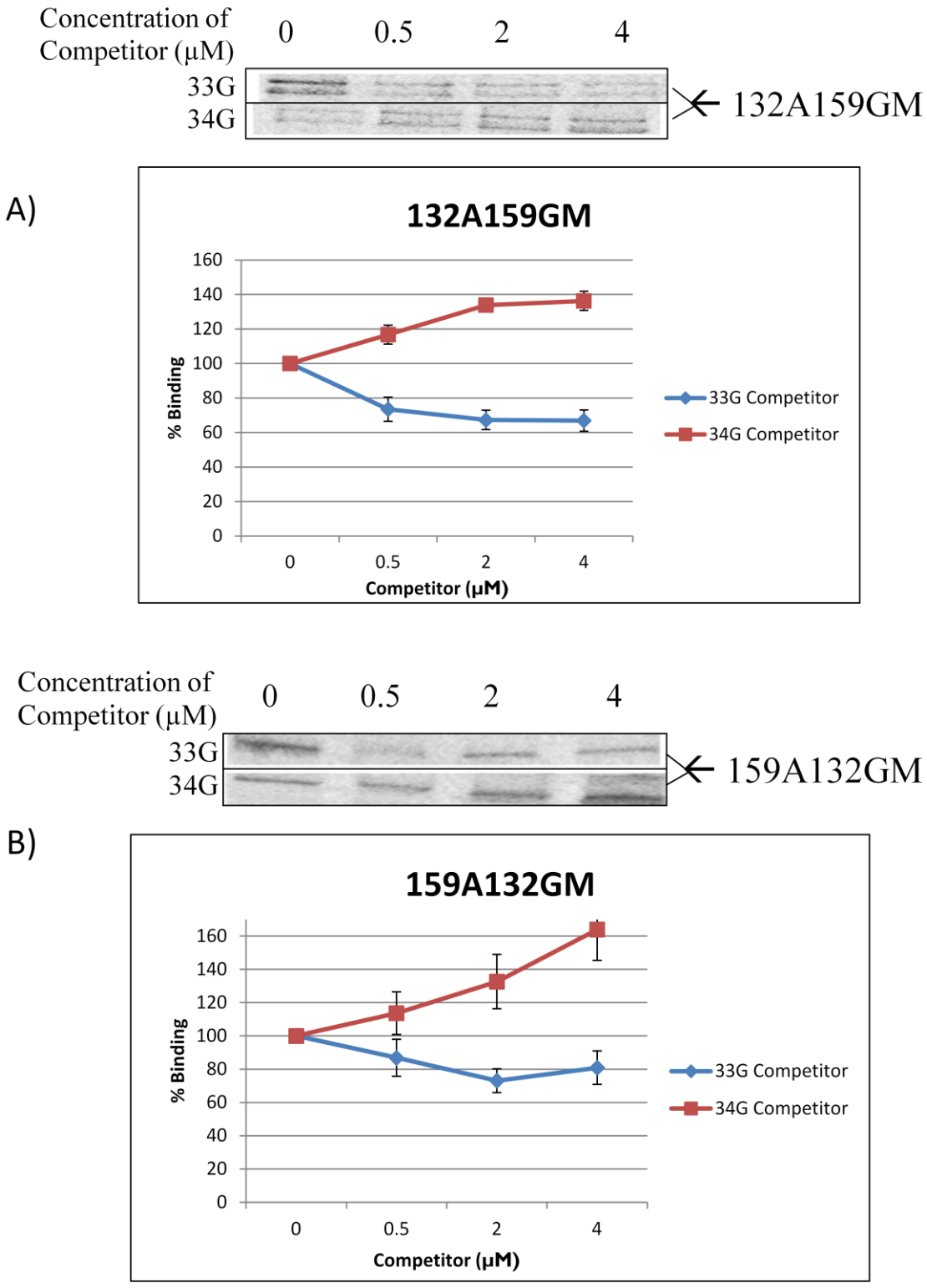


Figure 16 - *In vitro* Competitive Chloroplast Targeting Assays with A-Domain Swapped Proteins

Figure 16 - *In vitro* Competitive Chloroplast Targeting Assays with A-Domain Swapped Proteins. [^{35}S]-labelled 132A159GM (A) and 159A132GM (B) synthesized *in vitro* were targeted to isolated chloroplasts in the absence or presence of increasing concentrations of 33G or 34G as described in the Materials and Methods. Reactions were analyzed using SDS-PAGE, and targeting efficiency of 132A159GM (A) and 159A132GM (B) to wild type chloroplasts was analyzed by detecting radiolabelled proteins in dried gels using a phosphorimager. Efficiency of targeting is expressed as % Binding (percent of *in vitro* translated protein added to each reaction). Targeting efficiency is reflective of interaction with competitors. Gel images shown are representative of between 4-6 trials. Error bars represent standard error.

3.2 Solid-Phase Binding Assays

In an attempt to shed further light on the role played by the atToc132 A-domain in TOC complex assembly, *in vitro* solid-phase binding assays were performed. Interactions between 33G or 34G with atToc132, 132GM, 132A159GM and 159A132GM were compared by immobilizing either of 33G or 34G "bait" proteins onto Ni^{2+}-NTA resin via C-terminal-His_6 tags present on the proteins. A series of reactions were set up containing increasing amounts of the "bait" proteins (0, 100, 200, 300, 400, and 500 pmol) (Figures 17 and 18). In these assays, genuine interactions should result in the radiolabelled protein being pulled down via interaction with the immobilized binding partner, and can be quantified.

Data from individual experiments were compared based on a control lane included on the SDS-PAGE gel for each experiment, which contained the equivalent of 10% of the radiolabelled protein that was added to each reaction. Based on this, a value corresponding to 100% of added radiolabelled product could be determined using the Quantity One software (Bio-Rad). The efficiency of radiolabelled protein binding to the bait was determined for each amount of immobilized 33G/34G (100-500 pmol), after subtracting the background binding of radiolabelled protein to the resin in the absence of bait. Error bars represent standard error. Each data point represents the average of 3 trials.

The interaction between 34G and 132GM is markedly stronger than that between 33G and 132GM (Figures 17, 18). It is also noteworthy that 132GM is by far the strongest interacting partner of those tested for both 33G and 34G (Figures 17, 18). The interaction of full-length atToc132 (with A-domain present) is stronger with 34G than with 33G as the bait, which is in agreement with the hypothesis that the atToc132 A-

domain hinders interaction with atToc33.

The efficiency of the interactions between the A-domain swapped mutants (132A159GM and 159A132GM) with 33G was nearly identical. While it was expected that 132A159GM would interact more strongly with 34G than 33G (based on the atToc132 A-domain hypothesis), this may be an indication that the presence of the atToc159 GM-domains inherently favour interaction with 33G, despite the presence of the atToc132 A-domain. However, when interaction with 34G was tested, 132A159GM interacted more strongly than did 159A132GM (Figures 17, 18). This agrees with the hypothesis that the atToc132 A-domain hinders interaction with atToc33, as when it was present along with the 159 GM-domains (132A159GM), a preference can be observed for 34G when compared to 33G.

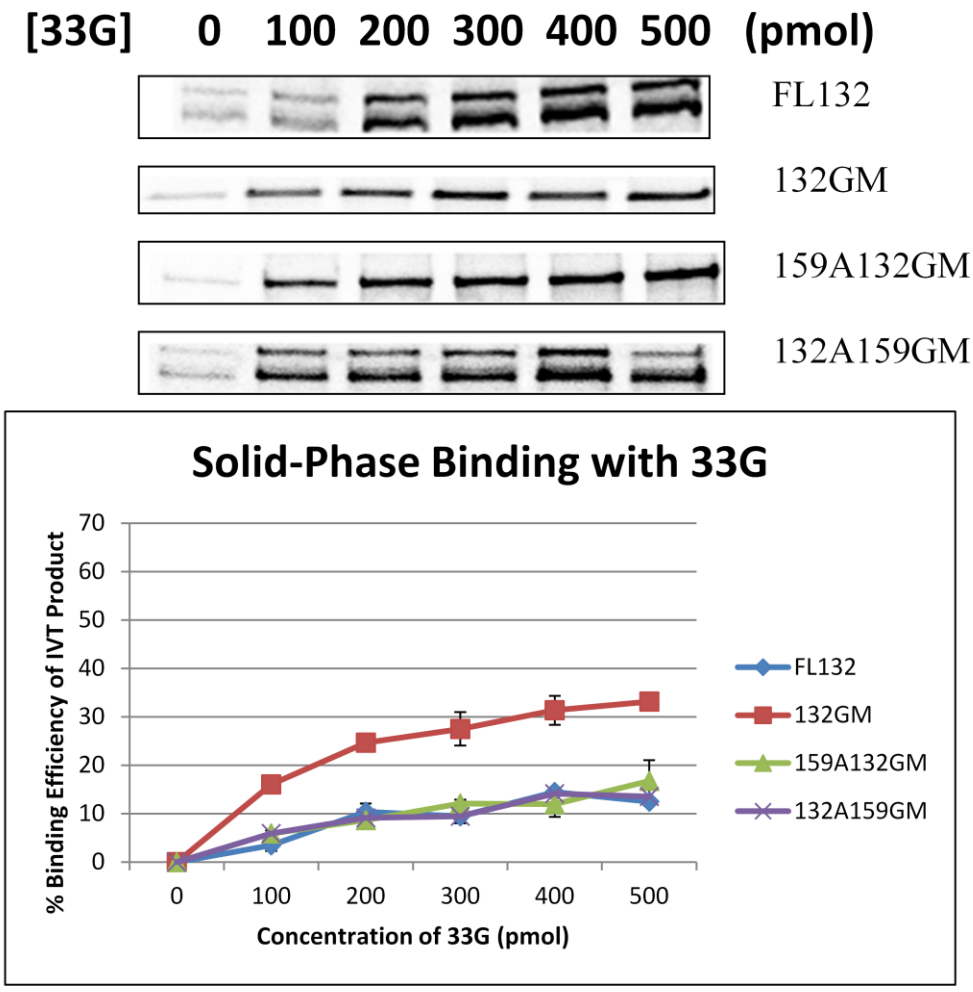


Figure 17 - Comparison of atToc132 (FL132), 132GM, 159A132GM and 132A159GM Interactions with 33G. *In vitro* solid-phase binding assay in the absence or presence of increasing amounts (100-500 pmol) of 33G as "bait" and [^{35}S]-labelled atToc132 (FL132), 132GM, 159A132GM or 132A159GM as "prey" in order to test and compare interaction preferences. Efficiency of binding is expressed as % Binding of IVT (*in vitro* translated) product added to each reaction. Binding efficiency is representative of interaction of the "bait" (33G) with the "prey" IVT protein. Gel images shown are representative of 3 trials. Error bars represent standard error.

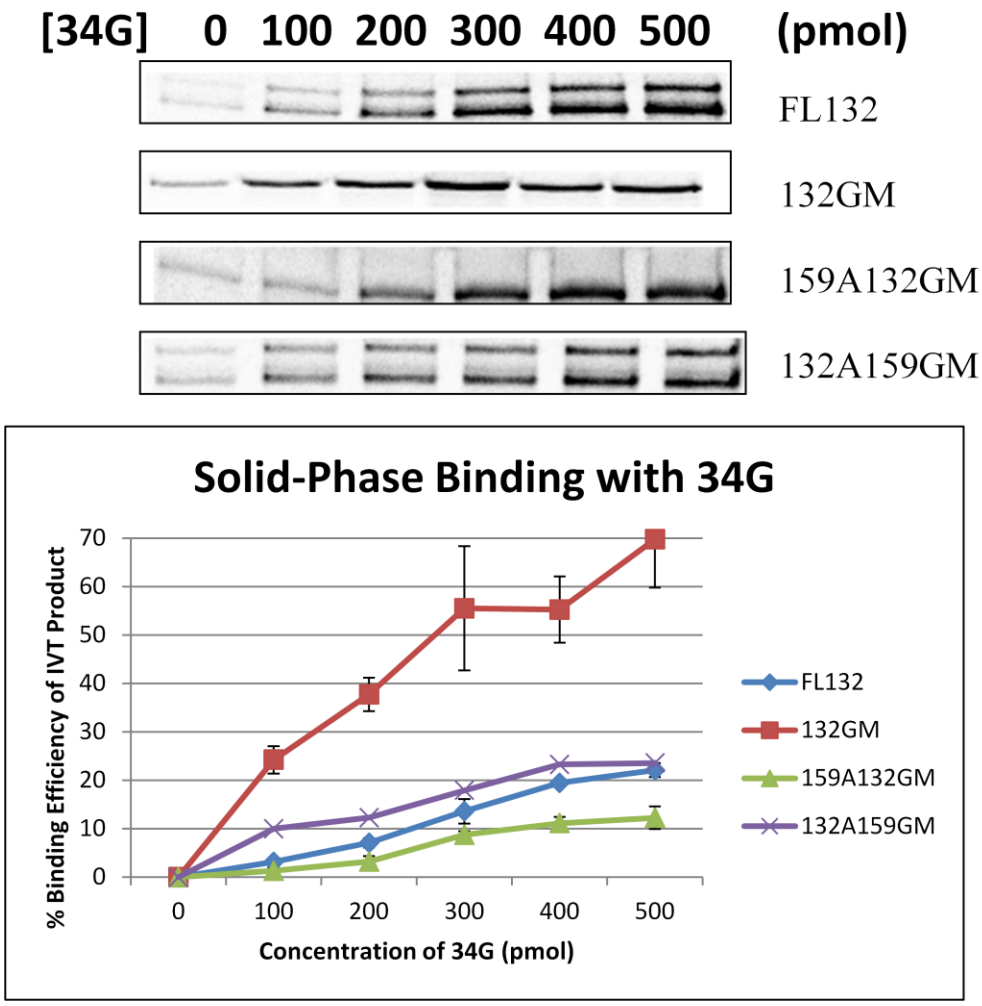


Figure 18 - Comparison of atToc132 (FL132), 132GM, 159A132GM and 132A159GM Interactions with 34G. *In vitro* solid-phase binding assay in the absence or presence of increasing amounts (100-500 pmol) of 33G as "bait" and [^{35}S]-labelled atToc132 (FL132), 132GM, 159A132GM or 132A159GM as "prey" in order to test and compare interaction preferences. Efficiency of binding is expressed as % Binding of IVT (*in vitro* translated) product added to each reaction. Binding efficiency is representative of interaction of the "bait" (34G) with the "prey" IVT protein. Gel images shown are representative of 3 trials. Error bars represent standard error.

4. Discussion

This project constitutes an investigation into the roles played by the TOC159 (atToc159, -132) and TOC34 (atToc33, -34) homologues from *Arabidopsis* in the assembly of structurally distinct TOC complexes. Formation of distinct complexes is thought to play an important functional role in plastid differentiation, as they confer the ability to balance import of different classes of preproteins that are required for the essential biochemical roles provided by plastids. By using *in vitro* chloroplast targeting assays and *in vitro* solid-phase binding assays, interactions between different isoforms of the TOC GTPase protein families were tested and compared in an effort to elucidate the interactions governing TOC complex formation. Experiments to test these interactions both in the context of binding with the chloroplast in targeting assays (where endogenous native TOC family members are present at the chloroplast surface), as well as in more chemically-defined systems once removed from the chloroplast altogether in solid-phase binding assays (only known amounts of proteins are present) were designed to show interaction preferences from different perspectives. In an effort to decipher the roles that individual protein domains play in this process, analysis was also extended to A-domain truncated mutants, and A-domain swapped mutants of some members of the TOC159 family from *Arabidopsis* to identify what role the A-domain plays versus the GTPase domain (and if this role is the same for the different isoforms tested of the TOC159 family).

The results present a challenge in deciphering which domains are responsible for conferring a preference for, and possibly inhibition against, interaction-specific binding partners. That the A-domains and GM-domains of atToc159 and atToc132 appear to have

different interaction preferences, makes it difficult to determine specifically how a given TOC complex is assembled. This challenge is made greater by the known differences of atToc33 vs. atToc34. As described in the introduction, the preference for atToc33 to be found in TOC complexes with atToc159, and for atToc34 to participate in TOC complexes with atToc132/120 is indicative of these proteins contributing to the differentiation of distinct targeting pathways to the chloroplast (Ivanova *et al.,* 2004). It has been demonstrated that atToc33 can be phosphorylated, while atToc34 cannot, and that the phosphorylation of atToc33 inhibits GTP binding, and as a consequence does not allow interaction with the precursor protein (Jelic *et al.,* 2003). It has also been shown that atToc33,-34 are found differentially in various tissue types (Gutensohn *et al.,* 2000). It has been suggested that atToc33 and -34 exhibit some preference for precursor binding, and that each is important for import of different subsets of proteins (Gutensohn *et al.,* 2000; Jelic *et al.,* 2003).

As illustrated in Figure 5, one model suggests that atToc33 homodimers (in the GDP loaded state) are disrupted by precursor binding, allowing one of the monomers to become active by binding GTP (Oreb *et al.,* 2011). It has been suggested that this step may be required for the complex to be capable of interaction with atToc159 (Sommer and Schlieff, 2009) (Figure 5a). Meanwhile, it has also been suggested that GTP could specifically disrupt atToc34-precursor interactions, perhaps encouraging association with atToc132/120 (Gutensohn *et al.,* 2000). If this phenomenon is more pronounced in the atToc132/120-atToc34 system than the atToc159-atToc33 system (or absent altogether in the latter), it could suggest one possible mechanism for regulating TOC complex formation. It is also accepted that heterodimerization between the members of the

Arabidopsis TOC34 family (atToc33, -34) also occurs (Kessler and Schnell, 2002), which is not addressed by the model described above. It could be that this GTP-driven mechanism is perhaps one of various regulatory systems influencing TOC complex assembly and thus regulating protein import.

Altogether, the results herein suggest that a role in TOC complex assembly is likely played by both the A-domains as well as the GM-domains of Toc159 isoforms, and perhaps a role is also played by an interaction preference on the part of atToc33 and atToc34 to interact with these different isoforms.

4.1 *In vitro* Targeting Assays

4.1.1 Full-Length and A-Domain Truncated Mutant Interactions with atToc33 and atToc34

Targeting of 159GM (A-domain deletion mutant) to chloroplasts is strongly inhibited by 33G (Figure 15a), indicating a strong interaction between these two proteins, whereas 34G does not inhibit the targeting of 159GM at all (Figure 15a). Targeting of 132GM is competed moderately by 34G, whereas 33G is not an effective competitor (Figure 15b). These data agree with the hypothesis that atToc159 interacts preferentially with atToc33 and atToc132 interacts preferentially with atToc34, although it also suggests that atToc132 does not have a strong preference for atToc34. Data presented in Figure 15 shows that at 4 μM of competitor, the resultant competition from 159GM's hypothesized preferred binding partner (33G) is approximately twice as strong as is that of 132GM's hypothesized preferred binding partner (34G) (binding efficiency: reduced to 78% for 132GM competed with 4 μM 34G, Figure 15b; reduced to 37% for 159GM

competed with 4 μM 33G, Figure 15a). This observation is in agreement with Ivanova *et al.* (2004), who observed a similar difference in binding strength between the two systems (interaction between atToc159 and atToc33 is much stronger than is that between atToc132/120 and atToc34) in a binding assay which compared the interactions. Furthermore, 33G is similarly effective at competing for the targeting of full-length atToc159, and the A-domain deletion mutant (159GM) to chloroplasts (Figures 14a, 15a). This could be an indication that atToc159 preference for atToc33 is not conferred by its A-domain, but rather the presence of the G- and -M domains is sufficient to confer specificity. This is consistent with the data shown in Figure 6 from Richardson (2008) which demonstrated a negligible effect on total chloroplast targeting when the A-domain of atToc159 was removed. It should also be noted that 33G is a more effective competitor of 159GM targeting (Figure 15a), than it is of 132A159GM (Figure 16b); binding efficiency was 37% vs. 66%, respectively, at 4 μM of competitor, which supports the hypothesis that 159GM is sufficient to confer specificity for 33G over 34G. It also supports the hypothesis that the 132A-domain inhibits an interaction with atToc33. This is because far less inhibition (indicating interaction) was observed for 132A159GM than 159GM alone when 33G was used as a competitor.

While it is hypothesized that for both atToc159 and atToc132 there is a dominant factor (e.g. A-domain) dictating the preferential interactions with other Toc GTPases (i.e. atToc33 and atToc34), the current data, as well as other recent data (e.g. Chen, 2011) suggests that these factors are not necessarily the same for atToc159 vs. atToc132. In other words, there may be some preference conferred from each of the A- and G-domains, and the mechanism may be different for each of the TOC159 isoforms.

4.1.2 A-Domain Swapped Mutant Interactions

Interestingly, targeting of both A-domain-swapped mutants to isolated chloroplasts was stimulated by the presence of 34G (Figure 16a and b). It must also be noted that 34G failed to cause any statistically relevant inhibition in all other targeting assays (atToc159 and atToc132, Figures 14a and b, respectively, as well as 159GM, Figures 15a) with the exception of 132GM (Figure 15b). This lack of competition observed by 34G and the stimulation of targeting in the case of the A-domain swapped mutants raises the possibility that somehow 34G did not only interact with the radiolabelled proteins, but that it somehow enhanced their interaction with the chloroplast surface. This was not expected because 33G/34G lack their transmembrane domains and therefore cannot insert into the membrane. However, as discussed in the introduction the G-domains have been shown previously to dimerize, and it is conceivable that 34G could interact with the radiolabelled proteins as well as with an endogenous atToc159/132/120 at the chloroplast surface. It seems reasonable to also consider the fact that atToc159 has been reportedly observed in a cytosolic form, as well as a membrane-associated form (Hiltbrunner *et al.*, 2001; Bauer *et al.*, 2002). While the existence of the soluble form of atToc159 is still a matter of debate (Soll and Schlieff, 2004), and it is unknown if the same can be said of atToc132, it could be a contributing factor to the data observed herein. It can be noted that a recent study using Toc159 and Toc132 from the single-cell C_4 system *Bienertia sinuspersici* (which share substantial identity to those found in *Arabidopsis*) observed both *B. sinuspersici* isoforms in membrane-associated as well as soluble forms (Lung and Chuong, 2012).

The stoichiometry of the Toc complex has not been determined precisely, but it

has also been reported that multiple copies of core TOC components have been found together (Agne and Kessler, 2009) (See Appendix regarding potential future experiment aimed at determining if multiple copies of atToc159 can be found in any given TOC complex). The observations described in this section together raise the possibility that interactions among TOC complex components may be more dynamic than previously thought. The potential existence of a soluble form of atToc159 makes it conceivable that atToc159 is capable of acting as a shuttle (Soll and Schlieff, 2004), delivering preproteins via interaction with a TOC complex that is formed in response to the arrival of atToc159 bearing its cargo. It is a possibility then, that the proposed shuttling effect could result in a 34G-(TOC159 homologue) heterodimer interaction with an endogenous TOC complex in the chloroplast targeting assay.

4.1.3 33G and 34G Interactions

A difference in the observed behaviours of 33G vs. 34G is shown in Figure 14. In targeting full-length atToc159 and atToc132 to chloroplasts using both 33G and 34G as competitors, it was found that 34G had a statistically insignificant effect on the targeting of both proteins (Figure 14a,b). On the other hand, 33G was an effective competitor of both atToc159 and atToc132.

These data do not agree with the hypothesis that atToc159 interacts preferentially with atToc33, while atToc132 interacts preferentially with atToc34. If that was the case, the expected observations would be that 33G would be a more effective competitor (as compared to 34G) for atToc159 targeting to chloroplasts, and that 34G would show greater competition (as compared to 33G) for targeting of atToc132. These unexpected

results may be related to the surprising observation that targeting of the A-domain swapped mutants was stimulated in the presence of 34G. The hypothesis of preferential interactions among TOC GTPases is partially based on differences between atToc33/34 (Jarvis *et al.,* 1998; Gutensohn *et al.,* 2000; Jelic *et al.,* 2003; Constan *et al.,* 2004; Ivanova *et al.,* 2004).

The observations that the two similar proteins atToc33 and atToc34 are both able to (at least partially) rescue *ppi*1 mutants (atToc33 deficient) (Jarvis *et al.,* 1998) and *ppi*3 mutants (atToc34 deficient), but with different phenotypic properties (Constan *et al.,* 2004), suggests the existence of at least two distinct forms of TOC complex. It was subsequently demonstrated that the two isoforms atToc33 and atToc34 could be found in distinct TOC complexes (Ivanova *et al.,* 2004). These observations suggest a probable difference in function between the two isoforms. The functional differences are reinforced by the fact that atToc33 and atToc34 are expressed differentially in various tissues (Jarvis *et al.,* 1998), and by the possibility that the two proteins exhibit some preference for precursor binding (Gutensohn *et al.,* 2000; Kubis *et al.,* 2003). Furthermore, it has also been reported that atToc33 is phosphorylated as a form of regulation, while atToc34 is not (Jelic *et al.,* 2003). It has been suggested that the atToc34 homodimer is disrupted by binding with a compatible protein precursor, allowing one monomer to exchange GDP for GTP (Gutensohn *et al.,* 2000) to disrupt atToc34-precursor binding, possibly promoting interaction with atToc132/120. According to this model, it may be easier to get atToc132/120 to interact with atToc34 than atToc159 to interact with atToc33 (Figure 5). If this is the case, a "GTP switch" (Bauer *et al.,* 2002) could have different effects on the two systems (specifically the

132/120-34 system would interact at a lower concentration of GTP than would the 159-33 system). Taken together, these data indicate that different amounts of each of atToc33 and atToc34 are not only necessary in tissues serving different functions (for example, photosynthetic tissues would require more of the isoform which has a preference for a photosynthetic subset of precursor proteins), but that they would likely also react differently to a change in GTP concentration. This GTP "switch" (Bauer *et al.,* 2002) might represent another form of control/response by the cell. If this form of cellular control affects the atToc33/34 isoforms differently, not only is their inherent preference for certain subsets of precursor proteins relevant, but also their preferences for the TOC159 isoforms (perhaps even a preference for particular domains of the TOC159 isoforms). All this taken with the fact that atToc159 and atToc132 also confer preferences for distinct subsets of precursors, and that interactions between the two TOC GTPase families are not exclusive (atToc159-atToc34 and atToc132/120-atToc33 interactions do occur, however not as frequently) presents a very intricate mechanism for regulation of TOC complex activity.

4.1.4 Summary of Targeting Observations

In summary, why 34G stimulates targeting to the chloroplast, and why this is only observed for the A-domain swapped mutants remains unknown. However, for atToc159 it appears that the G- and M-domains are sufficient to confer specificity of preferential interaction with atToc33, whereas for atToc132 it seems that the A-domain plays a more dominant role in influencing its interaction with the 33/34 family. This might in part explain the unexpected results when the A-domains are swapped. In other words, if the

A-domain confers specificity in one context (as is perhaps the case for atToc132, in conjunction with the other domains of that protein), whereas the G domain may play the dominant role in another case irrespective of the influence of the A-domain (as is perhaps the case for atToc159), it is quite reasonable to expect that swapping of the A-domains could lead to unexpected results. Additional evidence for this hypothesis is described in the next section, testing the interactions of the A-domain swapped mutants outside of the context of the chloroplast.

4.2 *In vitro* Solid-Phase Binding Assays

Previous members of the Smith lab have performed *in vitro* solid phase binding assays in an attempt to investigate the hypothesis that the atToc132 A-domain specifically hinders interaction with atToc33, thereby giving atToc132 a preference for atToc34. Specifically, binding assays were used to confirm the relative strength of interactions between atToc132 and 132GM with atToc33 and atToc34 (Chen, 2011; Ottaway, 2012). It was observed that 132GM binds with much higher efficiency to 33G than does full-length atToc132 (Figure 10, from Ottaway, 2012), presumably because the A-domain hinders the interaction. However, removal of the A-domain has less of an effect on the strength of interaction between atToc132 and atToc34G (Figure 11, from Ottaway, 2012). These data support the hypothesis that the atToc132 A-domain plays a role in preventing interaction with atToc33. This project not only involved repeating those experiments (due to an inadequate number of repeats in the previous studies), but extended the analysis to more fully address the potential role of the A-domain by also testing interactions between atToc33 and atToc34 with the A-domain swapped mutant

proteins 159A132GM and 132A159GM (Figures 17 and 18).

Previous findings (compare Figures 10 and 11, from Ottaway, 2012) found that removing the A-domain (deletion mutants) had less of an impact on the difference between observed interactions of atToc132 with 34G compared to 132GM with 34G than it did on the interactions between atToc132 with 33G compared to 132GM with 33G. In other words, it was observed that the difference between binding efficiencies when the A-domain was present (atToc132) and when it had been removed (132GM) was larger when the interaction was tested with 33G than with 34G. Contrary to this, the same effect was not observed in the current study (compare Figures 17 and 18). In fact, conversely, removal of the A-domain had a relatively much larger impact on the difference between the observed interactions between atToc132 with 34G compared to 132GM with 34G than the interactions between atToc132 with 33G compared to 132GM with 33G. It was also observed that the interaction between 34G and 132GM was relatively much stronger, as compared to that between 33G and 132GM (compare Figures 17 and 18). While the results seemingly disagree, it can be noted that the assays performed in the previous studies were performed again in this study due to large error bars (Figures 10 and 11, from Ottaway, 2012), which were significantly smaller when the assays were repeated (Figures 17 and 18). It is apparent that 132GM is the strongest interacting partner of those tested for both of 33G and 34G. Based on this observation, it would be expected that 132GM would be most strongly competed by each of 33G and 34G in the *in vitro* chloroplast targeting assays (because it demonstrates the strongest interaction of all proteins tested in the *in vitro* solid-phase binding assays); however, that was not the case (compare Figure 15b where little to no effect can be seen on targeting of 132GM to other

targeting data presented in Figures 14, 15, 16). This may be due to the fact that these TOC GTPase interactions are being tested out of the context of the chloroplast, with no associated endogenous protein interactions whatsoever, as would be present/occur in the chloroplast targeting assays and may be complicating the results.

Although the data herein contradicts one of the observations of Chen (2011) and Ottaway (2012), it is interesting that they do not refute the hypothesis that the atToc132 A-domain hinders interaction with atToc33. 132GM still shows stronger interaction with 33G than does full-length atToc132 (Figure 17), as would be expected according to the hypothesis. Meanwhile, comparison of Figures 17 and 18, reveals that atToc132 still interacts more strongly with 34G than 33G. It was also observed that 132A159GM interacted more strongly with 34G than 33G, which once again agrees with the 132 A-domain hypothesis (Figures 17 and 18). Taken together, none of these data refute the hypothesized function of the atToc132 A-domain in TOC complex assembly.

The efficiency of binding of the A-domain swapped mutants, 132A159GM and 159A132GM, with 33G was very similar (Figure 17). However, when tested with 34G, 132A159GM demonstrated much greater propensity to interact than did 159A132GM (Figure 18). While 132A159GM would be expected to interact more strongly with 33G than 159A132GM (due to a preference of 159GM for 33G), the observation that interactions were comparable for the two A-domain swapped mutants may indicate that the atToc159 A-domain is interacting with the 132 GM-domains in a way that it does not interact with the 159GM-domains. In other words, swapping the A-domains may be giving unexpected results because of intra-protein interactions which might not normally occur. The same can be said of the domain-swapped mutants with respect to interaction

with 34G. It would have been expected that 159A132GM would interact more strongly with 34G than did 132A159GM; however, once again there may be interactions occurring between the different domains of these TOC159 homologues which are leading to unexpected results.

While it would be seemingly useful to compare this result to that of the chloroplast targeting assays involving 132A159GM and 159A132GM (Figure16 a and b), this is made difficult because of the apparent stimulation of targeting that resulted from 34G. As explored earlier, there may be different interaction preferences exhibited by each domain of the different forms of the protein. Interestingly, 159A132GM demonstrates weak interaction with both 33G and 34G (Figures 17 and 18), while 132A159GM shows weak interaction with 33G, but a much stronger interaction with 34G. A direct comparison between the A-domain swapped proteins in the binding assay versus targeting assay is also made impractical once again because of the observation that targeting to the chloroplast is actually stimulated by 34G. While this effect could conceivably be due to interaction between the proteins (just as it is interaction being shown when a competitive effect is observed), it is not known how comparable these trends are to one another, as the exact mechanism of how stimulation would occur is not understood.

4.3 Conclusions

The major implications from this investigation are that the atToc132 A-Domain influences TOC complex assembly by specifically hindering interaction with atToc33, thereby giving it an apparent preference for atToc34. Meanwhile, in the case of atToc159

the presence of only the G- and M-domains is sufficient to confer a preference for interaction with atToc33. atToc159 appears to have a genuine preference for atToc33, with this preference seemingly working by a different mechanism than its homologue atToc132 which confers preference based on interaction with atToc33 being hindered. These interactions contribute to the formation of structurally and functionally distinct TOC complexes which have selectivity for different subsets of preproteins.

4.4 Future Experiments

An experiment that could directly add to the findings reported here would be to include atToc159 and 159GM in the *in vitro* solid-phase binding assays described in this project. These were not included among the assays as part of the current study because it was the atToc132 A-domain specifically which was being examined. However, in light of some of the conflicting results it may be useful to expand the experiments to include atToc159. It would be expected that atToc159 would have high binding efficiency with 33G, and low binding efficiency with 34G. It would also be expected based on previous observations that 159GM would have a very similar high binding efficiency with 33G (as it is hypothesized in this project that 159GM is sufficient to confer binding preference), and low binding efficiency with 34G.

Future work that could complement the findings here would extend analysis to atToc120 (including it in the *in vitro* targeting assays, and *in vitro* solid-phase binding assays). Being as the data here suggest that the contributing factors governing interaction preference within the TOC GTPases may be different for the isoforms examined (atToc159 and atToc132), it would be reasonable to extend the analysis to also include

atToc120. It would be expected to behave in a similar way as atToc132, however, as seen in this study there may be unexpected results if the A-domain is removed, or swapped with another isoform.

Another related experiment would be to further examine (i.e. confirm or disprove) the possibility that atToc159 occurs as both a cytosolic and membrane-associated form, as well as to extend the analysis to include the other isoforms atToc132 and atToc120. If this phenomenon is genuine, it could have important implications on understanding TOC complex formation and assembly.

In an attempt to bridge the gap between the two approaches used in this study (*in vitro* chloroplast targeting assays and *in vitro* solid-phase binding assays; the former testing protein interactions in the context of the chloroplast and the latter testing protein interactions outside of the context of the chloroplast), atToc33 or atToc34 together with atToc75 could be reconstituted into liposomes (Wallas *et al*., 2003), and targeting of atToc159, atToc132, and the domain-swapped versions of these proteins could be monitored in the absence or presence of increasing concentrations of GTP. These experiments would make it possible to test the hypothesis of a "GTP switch", and the possibility that the assembly of structurally distinct Toc complexes are differentially sensitive to GTP.

Recent work by Terry Lung has shown evidence that the transit peptide of an imported chloroplastic preprotein may exist on the C-terminal end of the protein (Lung and Chuong, 2012). This line of research should be pursued, as it could change the current understanding of plastid protein import.

The Appendix describes an investigation originating from this project that was not

completed, which was aimed at determining if more than one copy of atToc159 is present in a given TOC complex (See Appendix).

4.5 Implications and Integration of This Research In A Broader Biological Context

The ability for the chloroplast to import proteins, and be able to regulate said import, is crucial for plant growth and development. Protein trafficking is an important process for all eukaryotes, and fully understanding model systems (such as chloroplast protein import in *Arabidopsis*) contributes to a general understanding of the function and regulation of similar processes in nature.

Plants are an essential part of the global ecosystem, and not only produce the oxygen that animals need to breathe, but are an essential food source as well. As the global food demand rises and new crop-growing strategies are developed, a full understanding of the processes within the plant cell will be extremely useful, particularly in the context of creating new transgenic crops. While the research here is one small piece of a very large puzzle, it takes contributions from many different fields of plant biology to have an impact on problems faced currently and problems that will arise in the future.

5. References

Agne, B. and Kessler, F. (2009). Protein transport in organelles: The Toc complex way of preprotein import. *The FEBS Journal, 276,* 1156-1165.

Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature, 408 (6814),* 796-815.

Aronsson, H. and Jarvis, P. (2009). The chloroplast protein import apparatus, its components, and their roles. *Plant Cell Monographs, 13,* 89-123.

Aronsson, H. and Jarvis, P. (2011). Dimerization of TOC receptor GTPases and its implementation for the control of protein import into chloroplasts. *Biochemical Journal, 436,* e1-e2.

Azpiroz-Leehan, R. and Feldmann, K. A. (1997). T-DNA insertion mutagenesis in *Arabidopsis*: going back and forth. *Trends in Genetics, 13 (4),* 152-156.

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

Bauer, J., Hiltbrunner, A. and Kessler, F. (2001). Molecular biology of chloroplast biogenesis: gene expression, protein import and intraorganellar sorting. *Cellular and Molecular Life Sciences, 58,* 420-433.

Bauer, J., Hiltbrunner, A., Weibel, P., Vidi, P-A., Alvarez-Huerta, M., Smith, M. D., Schnell, D. J. and Kessler, F. (2002). Essential role of the G-domain in targeting of the protein import receptor atToc159 to the chloroplast outer membrane. *The Journal of Cell Biology, 159 (5),* 845-854.

Becker, T., Jelic, M., Vojta, A., Radunz, A., Soll, J. and Schlieff, E. (2004). Preprotein recognition by the Toc complex. *The European Molecular Biology Organization Journal, 23,* 520-530.

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

Chen, K., Chen, X. and Schnell, D. J. (2000). Initial Binding of Preproteins Involving the Toc159 Receptor Can Be Bypassed during Protein Import into Chloroplasts. *Plant Physiology, 122,* 813-822.

Chen, Y. (2011). The role of the acidic domain of the Toc159 protein import receptor family in Toc complex assembly in *Arabidopsis thaliana*. A thesis for the degree of Master of Science, Wilfrid Laurier University.

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

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

Dhanoa, P. K., Richardson, L. G. L., Smith, M. D., Gidda, S. K., Henderson, M. P. A., Andrews, D. W. and Mullen, R. T. (2010). Distinct Pathways Mediate the Sorting of Tail-Anchored Proteins to the Plastid Outer Envelope. *PLoS ONE 5(4): e10098.* doi:10.1371/journal.pone.0010098

Gutensohn, M., Schulz, B., Nicolay, P. and Flugge, U. (2000). Functional analysis of the two *Arabidopsis* homologues of Toc34, as a component of the chloroplast protein import apparatus. *The Plant Journal 23 (6),* 771-783.

Hiltbrunner, A., Bauer, J., Vidi, P-A., Infanger, S., Weibel, P., Hohwy, M. and Kessler, F. (2001). Targeting of an abundant cytosolic form of the protein import receptor atToc159 to the outer chloroplast membrane. *The Journal of Cell Biology, 154 (2),* 309-316.

Inoue, H., Rounds, C. and Schnell, D. J. (2010). The Molecular Basis for Distinct Pathways for Protein Import into *Arabidopsis* Chloroplasts. *The Plant Cell, 22,* 1947-1960.

Inoue, K. (2011). Emerging roles of the chloroplast outer envelope membrane. *Trends in Plant Science,* doi:10.1016/j.tplants.2011.06.005.

Ivanova, Y., Smith, M. D., Chen, K. and Schnell, D. J. (2004). Members of the Toc159 Import Receptor Family Represent Distinct Pathways for Protein Targeting to Plastids. *Molecular Biology of the Cell, 15,* 3379-3392.

Jarvis, P., Chen, L-J., Li, H-M., Peto, C. A., Fankhauser, C. and Chory, J. (1998). An *Arabidopsis* Mutant Defective in the Plastid General Protein Import Apparatus. *Science, 282,* 100-103.

Jelic, M., Soll, J. and Schlieff, E. (2003). Two Toc34 Homologues with Different Properties. *Biochemistry, 42,* 5906-5916.

Keegstra, K. and Cline, K. (1999). Protein Import and Routing Systems of Chloroplasts. *The Plant Cell*, *11*, 557-570.

Kessler, F. and Schnell, D. (2002). A GTPase Gate for Protein Import into Chloroplasts. *Nature Structural Biology, 9 (2)*, 81-83.

Kikuchi, S., Hirohashi, T. and Nakai, M. (2006). Characterization of the Preprotein Translocon at the Outer Envelope Membrane of Chloroplasts by Blue Native PAGE. *Plant Cell Physiology, 47 (3),* 363-371.

Koenig, P., Oreb, M., Rippe, K., Muhle-Goll, C., Sinning, I., Schlieff, E. and Tews, I. (2008). On the Significance of Toc-GTPase Homodimers. *Journal of Biological Chemistry*, *238* (34), 23104-23112.

Kouranov, A. and Schnell, D. J. (1997). Analysis of the Interactions of Preproteins with the Import Machinery over the Course of Protein Import into Chloroplasts. *The Journal of Biological Chemistry, 139 (7),* 1677-1685.

Kubis, S., Baldwin, A., Patel, R., Razzaq, A., Dupree, P., Lilley, K., Kurth, J., Leister, D. and Jarvis, P. (2003). The Arabidopsis *ppi*1 Mutant Is Specifically Defective in the Expression, Chloroplast Import, and Accumulation of Photosynthetic Proteins. *The Plant Cell*, *15,* 1859-1871.

Kubis, S., Patel, R., Combe, J., Bedard, J., Kovacheva, S., Lilley, K., Biehl, A., Leister, D., Rios, G., Koncz, C. and Jarvis, P. (2004). Functional Specialization amongst the Arabidopsis TOC159 Family of Chloroplast Protein Import Receptors. *The Plant Cell*, *16*, 2059-2077.

Lee, D. W., Lee, S., Oh, Y. J. and Hwang, I. (2009). Multiple Sequence Motifs in the Rubisco Small Subunit Transit Peptide Independently Contribute to Toc159-Dependent Import of Proteins into Chloroplasts. *Plant Physiology*, *151,* 129-141.

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

Lung, S. and Chuong, S. D. X. (2012). A Transit Peptide-Like Sorting Signal at the C Terminus Directs the *Bienertia sinuspersici* Preprotein Receptor Toc159 to the Chloroplast Outer Membrane. *The Plant Cell*, *24*, 1560-1578.

Meinke, D., Cherry, J., Dean, C., Rounsley, S. and Koornneef, M. (1998). *Arabidopsis thaliana*: A Model Plant for Genome Analysis. *Science, 282 (5389),* 662-682.

Nelson, D. L. and Cox, M. M. Lehninger Principles of Biochemistry 4th Ed. *W.H. Freeman and Company.* (2005)

Oreb, M., Hofle, A., Koenig, P., Sommer, M. S., Sinning, I., Wang, F., Tews, I., Schnell, D. J. and Schlieff, E. (2011). Substrate Binding Disrupts Dimerization and Induces Nucleotide Exchange of the Chloroplast GTPase TOC33. *Biochemical Journal, 436 (2),* 313-319.

Ottaway, J. (2012). The role of the acidic domain of Toc132 in interactions with the Toc34 family proteins in *Arabidopsis thaliana*. A thesis for the degree of Honours Bachelor of Science, Wilfrid Laurier University.

Richardson, L. (2008). Structure-function analysis of the acidic domain of the Arabidopsis Toc159 receptors. A thesis for the degree of Master of Science, University of Waterloo.

Richardson, L., Jelokhani-Niaraki, M. and Smith, M. D. (2009). The acidic domains of the Toc159 chloroplast preprotein receptor family are intrinsically disordered protein domains. *BioMed Central Biochemistry*, 10:25.

Richter, S., Zhong, R. and Lamppa, G. (2005). Function of the stromal processing peptidase in the chloroplast import pathway. *Physiologia Plantarum, 123,* 362-368.

Rounds, C. M. Assembly of structurally and functionally distinct protein import translocons in chloroplasts. *University of Massachusetts Doctor of Philosophy Thesis Dissertation* (2007).

Salinas, J. and Sanchez-Serrano, J. *Arabidopsis* Protocols 2nd Ed. *Humana Press.* (2006).

Schleiff, E., Soll, J, Küchler, M., Kühlbrandt, W. and Harrer, R. (2003). Characterization of the translocon of the outer envelope of chloroplasts. *The Journal of Cell Biology, 160 (4),* 541-551.

Schnell, D. J., Blobel, G., Keegstra, K., Kessler, F., Ko, K. and Soll, J. (1997). A consensus nomenclature for the protein-import components of the chloroplast envelope. *Trends in Cell Biology, 7 (8),* 303-304.

Smith, M.D., Fiuzpatrick, L., Keegstra, K.. and Schnell, D.J. (2002a). *In vitro* analysis of chloroplast protein import. *Current Protocols in Cell Biology.* Eds. Bonifacina JS, Dasso M, Harford JB, Lippincott-Schwartz J, Yanada KM. New York: John Wiley and Sons Inc. 11.16.1-11.16.21.

Smith, M. D., Hiltbrunner, A., Kessler, F. and Schnell, D. J. (2002b). The targeting of the atToc159 preprotein receptor to the chloroplast outer membrane is mediated by its GTPase domain and is regulated by GTP. *The Journal of Cell Biology*,

159 (5), 833-843.

Smith, M. D., Rounds, C. M., Wang, F., Chen, K., Afitlhile, M. and Schnell, D. (2004). atToc159 is a selective transit peptide receptor for the import of nucleus-encoded chloroplast proteins. *The Journal of Cell Biology, 165 (3),* 323-334.

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

Soll, J. and Schlieff, E. (2004). Protein import into chloroplasts. *Nature Reviews. Molecular Cell Biology, 5 (3),* 198-208.

Sommer, M. S. and Schlieff, E. (2009). Molecular interactions within the plant TOC complex. *The Journal of Biological Chemistry, 390,* 739-744.

Sun, Y. J., Forouhar, F., Li Hm, H. M., Yeh, Y. H., Kao, S., Shr, H. L., Chou, C. C., Chen, C. and Hsiao, C. D. (2002). Crystal structure of pea Toc34, a novel GTPase of the chloroplast protein translocon. *Nature Structural Biology, 9 (2),* 95-100.

Wallas, T.R., Smith, M. D., Sanchez-Nieto, S. and Schnell, D. J. (2003). The Roles of Toc34 and Toc75 in Targeting the Toc159 Preprotein Receptor to Chloroplasts. *The Journal of Biological Chemistry, 278 (45),* 44289-44297.

Weston, K. (2011). A funtional characterization of the chloroplast protein import receptor Toc159. A thesis for the degree of Master of Science, Wilfrid Laurier University.

Wise, R. R. and Hoober, J. K. Advances in Photosynthesis and Respiration, Vol. 23: The Structure and Function of Plastids. *Springer*. (2006).

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

6. Appendix

An initial objective of this project was to attempt to answer the fundamental question: can more than one copy of atToc159 be found in a given TOC complex? A number of ratios for the core components of the TOC complex have been reported, including 1 : 4 : 4-5 (Schleiff *et al.*, 2003) and 1 : 3 : 3 (Kikuchi *et al.*, 2006) for TOC159 : TOC75 : TOC34, respectively. While ratios for the core TOC proteins in a complex have been proposed, nobody has quantified how many copies of the core TOC proteins are present in a given complex (Rounds, 2007). It has also been observed that atToc132 and atToc120 can be found together in an atTOC complex (Ivanova *et al.*, 2004). Using a transgenic *Arabidopsis* strain, it may be possible to determine if only one copy of TOC159 can be found in a TOC complex, or if more than one copy of the protein can be found in a given complex.

Rationale

The transgenic strain of *Arabidopsis thaliana* that may make this possible was made by Caleb Rounds during the course of his research at the University of Massachusetts. It is a cross of two *ppi2* mutant plants (deficient in atToc159) (Bauer *et al.*, 2000), each of which was transformed with a different epitope-tagged version of the atToc159 protein (-His_6 tag and -Myc tag). The two tagged plant lines were then crossed, and the tagged mutants (the "Myc x His" plant line) were selected for using the herbicide glufosinate ammonium (BASTA) which the mutants are resistant to because of a resistance gene included with the vector encoding the epitope-tagged versions of the proteins (Rounds, 2007). The fact that this plant line produces two epitope-tagged versions of atToc159 provides an opportunity to examine its stoichiometry within an

atToc complex.

While it will not be possible to quantify the proteins of a given TOC complex with the tools available, working with a transgenic version of *Arabidopsis* would allow one to ask if there are multiple copies of atToc159 in a given atTOC complex, or just one.

With the Myc x His plant line simultaneously producing atTOC159 with two different epitope tags, it provides a unique opportunity to examine whether or not more than one copy of atToc159 associates in the same complex. Even if two copies of atToc159 are not in direct contact but are both present in a given TOC complex, some of these complexes will have a copy of atToc159 with each affinity tag. The way this plant line can be used to examine how many copies of atTOC159 are present is by exploiting antibodies against the epitope tags, and utilizing two different techniques: co-immunoprecipitation and Western blotting.

Using co-immunoprecipitation, an entire atTOC complex can be pulled specifically out of a solubilized chloroplast homogenate by using antibodies to target one particular protein (Ivanova *et al.,* 2004). In this case the target will be an epitope tag fused to atToc159. Appropriate antibodies were obtained from Millipore (anti-His monoclonal mouse Ab, clone HIS.H8, Cat.# 05-949; anti-Myc monoclonal mouse Ab, clone 9E10, Cat.# 05-419).

By immobilizing the antibody against either one of the epitope tags present on Toc159 to a substrate, and incubating with solubilized chloroplast envelope membranes, the affinity-tagged protein will be immobilized along with the rest of the complex it is associated with. Once an atTOC complex has been pulled out of solution by an antibody against one of the epitope tags present on atToc159, it can be eluted from the substrate

that the antibody is fixed to. The purified protein mixture can then be precipitated from the eluate in order to concentrate it and subsequently dissolved in SDS-PAGE sample buffer. The proteins of the complex can then be separated on an SDS-PAGE gel, transferred onto a stable nitrocellulose membrane via Western blot transfer and probed with the antibody for the other affinity tag (Ivanova *et al.*, 2004). If a signal is obtained, it would mean that the copy of atToc159 with the second affinity tag was originally present when the complex was pulled out of solution. This would demonstrate that at least two copies of the protein were present in one atTOC complex at the same time. To confirm that entire atTOC complexes are being isolated, the immunoprecipitate can be probed with antibodies against the other core TOC proteins, atToc75 and atToc33/34.

I was successful in showing (via Western blot) that both epitope-tagged versions of atToc159 were in fact being produced in the transgenic plants, and were compatible with the anti-myc and anti-his antibodies (Figure 19). The Western blot was performed on the total insoluble chloroplast fraction (which includes the chloroplast outer membrane).

Immunoprecipitation reactions were then carried out as the next step of this experiment, once again on total chloroplast insoluble fraction, using anti-His mAb Magbeads (Genscript Cat.# L0025). After elution of the protein, the eluate, was precipitated due to the volume being too large to load onto an SDS-PAGE gel (a number of different precipitation methods were compared), and Western blot analysis was performed on the eluate, the supernatant and the starting material. A signal can be detected for Toc159 when probed with an anti-myc antibody, however is absent when probed with an anti-his antibody (Figure 20). Another problem arises from the fact that if

a signal is detectable in the supernatant, it should be detectable in the starting material as well. The detection of the myc-tagged protein, but not the his-tagged version led to the conclusion that the epitope-tagged versions of atToc159 are below the detection limit (as the starting material is only 10% of what was used in the immunoprecipitation reaction). Being as the transgenic Myc x His line of *Arabidopsis* being used for this investigation was selected for, but not screened, not every seedling will be producing *both* tagged versions, resulting in sporadic expression levels and inconsistent results.

The reason this initial primary objective was not completed is because while the transgenic *Arabidopsis* line had been put through a selection process to identify plants which had been transformed, a line was not established in which it was known that *both* epitope-tagged versions of atToc159 were being produced at high levels in the same plant. In order to identify and establish a line in which both versions of the tagged atToc159 proteins are present, a PCR screen was initiated. 40 of the transgenic seeds from Caleb Rounds (which had been selected for, but not screened) were sown, and PCR primers were designed to differentiate between and detect each tagged version of atToc159. The relatively simple PCR reactions only needed to show amplification for each tag, and reactions appeared to be successful. Tissue was taken from each seedling and the DNA was extracted using a DNA extraction kit (Qiagen DNeasy Plant kit, Cat.# 69104). After quantification, the DNA from each tissue sample was used in 2 different PCR reactions: one to test for the presence of each epitope tag. A positive test for both tags would mean that seeds from that particular plant were harvested and re-sown, to have their progeny undergo the same PCR screen. This process was to be repeated until such time as several generations of progeny continue to show the presence of both tagged

proteins, thus establishing a reliable (pure) line. Of the initial 40 seeds sown, 10 of the resultant seedlings showed the presence of both epitope tags. Of their progeny, only 3 seedlings showed the presence of both tags. It was realized at this point that the initial screen should have included a much larger number of seeds, and of the seedlings which showed the presence of both tags, many of their progeny should have been tested instead of only a couple. Due to time constraints this objective was dropped from this project, however using what has been accomplished so far can be used as a starting point for continuing the study in the future.

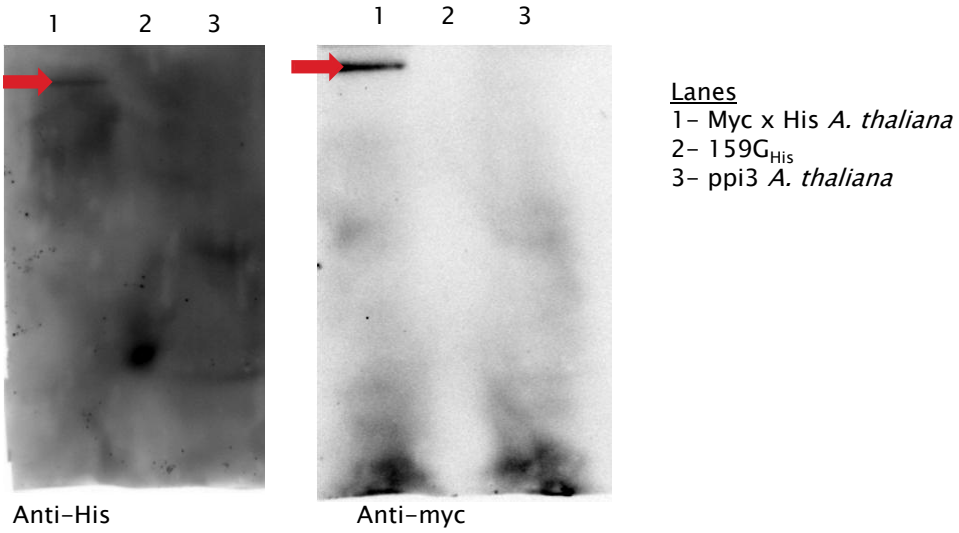
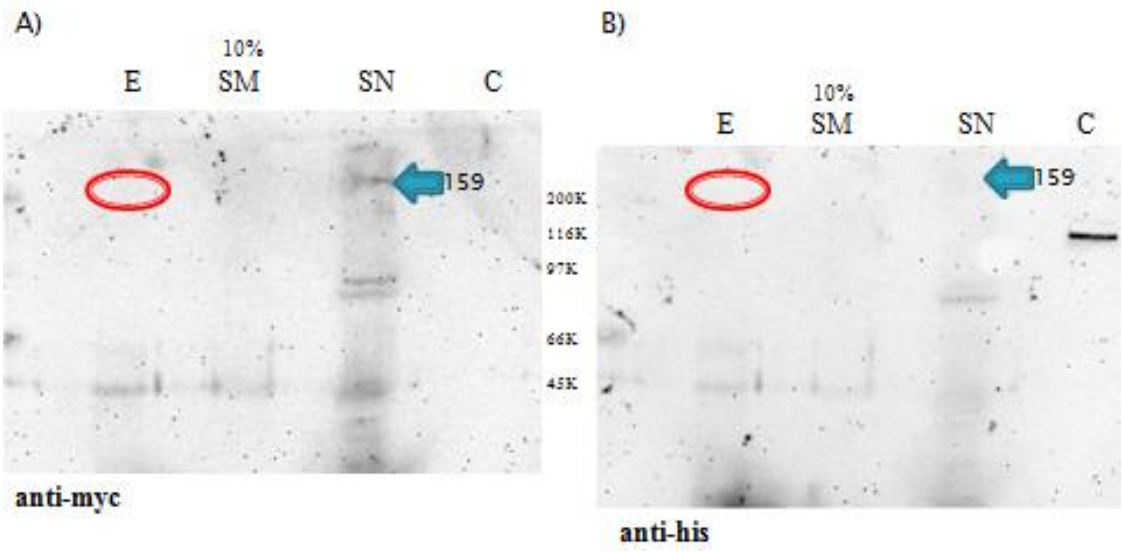


Figure 19 - Western Blots of Total Chloroplast Insoluble Fraction (including solubilized membranes) to Test Antibodies. Blots were probed with either anti-his or anti-myc antibodies. These images show that both tagged forms of atToc159 are being produced in the transgenic *Arabidopsis* strain, and that the antibodies being used are compatible.



E- Eluate (Immunoprecipitated Proteins)
SM- Starting material
SN- Supernatant (flowthrough)
C- Control ($159A_{his}$)

Figure 20 - Western Blot of Immunoprecipitated Proteins. Western Blots which were performed with both (A) anti-myc and (B) anti-his antibodies. Starting material used was total solubilized chloroplast insoluble fraction (including solubilized outer membrane). Eluate represents proteins eluted from anti-his mAb Magbeads which were precipitated then dissolved in 2xSDS-Sample Buffer. Starting material represents 10% of the starting material used in the immunoprecipitation. Supernatant represents the supernatant from the immunoprecipitation. Control represents the 159 A-domain with a -his tag. A signal can be seen in the supernatant for Toc159 when probed with an anti-myc antibody, however is absent when probed with an anti-his antibody. It is hypothesized that the epitope-tagged versions of atToc159 are below the detection limit of the antibodies.

Primer Design

Forward Primer (shared for both tags) and -Myc Reverse Primer

Both tags are located on the C-terminal end of the protein, meaning that one forward (sense) primer could be used to test for both tags, while two reverse (antisense) primers were needed (one for each tag).

Primers were designed based on sequences provided by Caleb Rounds, and are shown here. The forward primer (which was used in testing for the presence of both -his and -myc tags) is at the top of the following sequence, and is in bold and underlined (5'-GCGTCAATGCAGAACACAGTCT-3'). The sequence underlined at the bottom is that used to generate the antisense primer for testing for the -myc tag (actual antisense primer ordered: 5'-TCTTCAGAAATAAGTTTTTGTTCGTC-3'), while the -myc tag itself is highlighted.

5'...GTGGATATGACG**GCGTCAATGCAGAACACAGTCT**TGCTCTAGCTAGCCG

GTTCCCTGCCACAGCTACTGTCCAAGTCACCAAGGACAAGAAAGAGTTCAAC

ATTCATCTGGACTCCTCTGTGTCTGCTAAGCACGGGGAGAATGGATCCACCAT

GGCAGGGTTCGATATTCAGAATGTA[GGCAAGCAGCTGGCATATGTGGTCAG

AGGAGAAACCAAATTCAAGAATTTGAGGAAGAACAAGACAACTGTTGGAGG

GTCAGTGACATTCTTGGGAGAGAACATCGCCACTGGGGTCAAACTCGAGGAC

CAAATAGCACTGGGGAAAAGGTTGGTGCTTGTGGGCAGCACTGGGACAATGC

GATCACAGGGAGATTCGGCCTATGGTGCGAACCTCGAGGTCAGGCTTAGGGA

AGCTGATTTCCCAATTGGACAGGACCAATCTTCTTTTGGGCTGTCTCTGGTAA

AGTGGAGAGGCGATTTAGCCCTTGGAGCCAATCTCCAATCTCAAGTCTCTGTT

GGAAGGAACTCAAAGATTGCGCTTCGTGCAGGACTTAACAACAAGATGAGCG

GACAGATCACAGTCAGAACCAGCAGCTCGGATCAGTTGCAAATCGCTCTCAC
AGCCATTCTTCCAATTGCCATGTCCATCTACAAGAGCATTCGACCCGAAGCG
ACGAACGACAAGTACAGCATGTACGTCGACGAACAAAAACTTATTTCTGAAG
AAGATCTGATCCTCTAGAGCTCGGTACCAAGC - 3'

-His Reverse Primer

The underlined sequence is that used to generate the antisense primer for testing for the -his tag (actual antisense primer ordered: 5'-TGGTGGTGGTGGTCGCTGAGTG-3'), while the -his tag itself is highlighted.

5'...AGCTTGCGGCCGCACTCGAGCACCACCACCACCACCACTGAGATCCGGCT
GCTAACAA - 3'