

Wilfrid Laurier University

Scholars Commons @ Laurier

---

Theses and Dissertations (Comprehensive)

---

2009

## Characterization of the Interaction between Arabidopsis Toc159 and Toc33 GTPase Domains

Wesley A. Farquharson  
*Wilfrid Laurier University*

Follow this and additional works at: <https://scholars.wlu.ca/etd>



Part of the [Biology Commons](#)

---

### Recommended Citation

Farquharson, Wesley A., "Characterization of the Interaction between Arabidopsis Toc159 and Toc33 GTPase Domains" (2009). *Theses and Dissertations (Comprehensive)*. 964.  
<https://scholars.wlu.ca/etd/964>

This Thesis is brought to you for free and open access by Scholars Commons @ Laurier. It has been accepted for inclusion in Theses and Dissertations (Comprehensive) by an authorized administrator of Scholars Commons @ Laurier. For more information, please contact [scholarscommons@wlu.ca](mailto:scholarscommons@wlu.ca).



Library and Archives  
Canada

Published Heritage  
Branch

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

Bibliothèque et  
Archives Canada

Direction du  
Patrimoine de l'édition

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file* *Votre référence*  
*ISBN: 978-0-494-64354-9*  
*Our file* *Notre référence*  
*ISBN: 978-0-494-64354-9*

**NOTICE:**

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

---

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

**AVIS:**

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

---

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

  
**Canada**



**Characterization of the interaction between Arabidopsis Toc159 and Toc33 GTPase  
domains**

by

Wesley A. Farquharson

BSc, Wilfrid Laurier University, 2006

THESIS

Submitted to the Biology Department of Wilfrid Laurier University

In partial fulfillment of the requirements for

Master of Science

in

Integrative Biology

Wilfrid Laurier University

2009

Wes Farquharson © 2009

## Abstract

Protein-protein interactions are critical for many biochemical processes and for cellular function. Understanding interactions between proteins and ligands, as well as other proteins are necessary for the modelling of biological systems. Protein import into chloroplasts is one such system where the understanding of interactions between members of the translocation complex is limited. The translocon at the outer envelope membrane of chloroplasts (Toc complex) consists of three core components: Toc75, Toc159, and Toc34. Members of the Toc159 and Toc34 gene families in *Arabidopsis thaliana* have been shown to assemble differentially into structurally and functionally distinct complexes. More specifically, there is evidence that complexes containing atToc159 and atToc33 (the prefix “at” identifies the species of origin, *Arabidopsis thaliana*) import photosynthetic proteins, whereas complexes containing atToc132/120 and atToc34 seem to preferentially import non-photosynthetic proteins. The interactions between atToc159 and atToc33 or atToc132 and atToc34 are mediated by their GTPase domains. The current research used recombinant GTPase domains to investigate the specifics of the interaction between atToc159G and atToc33G in more detail using molecular and biophysical approaches. Methods included using blue-native PAGE, fluorescence spectroscopy and circular dichroism spectroscopy to characterize the interaction between atToc159G and atToc33G. Objectives were to characterize the interaction between atToc159G and atToc33G as well as attempt to manipulate the monomer:dimer equilibrium for individual proteins for the purpose of studying heterodimer formation. A secondary objective was to generate single tryptophan mutants of atToc159G (W973F and W1056F) to more precisely characterize the interaction between proteins. Results indicate that only a small percentage of the total protein exists as heterodimers, making quantification using fluorescence spectroscopy and circular dichroism spectroscopy difficult. Nonetheless, it was demonstrated that heterodimer formation occurs between the G-domains of atToc159 and atToc33. The data also suggest that higher order oligomers may form, which may be indicative of multiple interaction domains among proteins as previously suggested by stoichiometric studies on isolated Toc complexes. The effects of chelating agents, such as EDTA and

Chelex, produced significant structural changes which also inhibited homo-dimerization indicating the specific need for GTP-nucleotide to maintain a functional conformation. Single tryptophan mutants of atToc159G also produced major structural changes to recombinant proteins highlighting the importance of these residues for overall structure of the protein.

## **Acknowledgements**

Firstly, I would like to thank my supervisor Dr. Matthew Smith for his expertise, encouragement and support in both my academic and personal life during this project. I would also like to thank Dr. Arthur Szabo for his expertise in fluorescence spectroscopy and the time he spent to explain the concepts of fluorescence to me. Thanks to Dr. Masoud Jelokhani-Niaraki for his expertise in both fluorescence and circular dichroism spectroscopy and his continual support and interest in my project. Thanks to all my lab mates and support staff at Laurier over the past two years who have made my experience one to remember: Kyle Weston, Spence MacDonald, Patrick Hoang, Marina Ivanova, Sid Dutta, Mike Pyne, Fern McSorley, as well as Gena Braun and Lynn Richardson. I would like to pay special thanks to both Gena Braun and Lynn Richardson for their help with training and continued support on instruments and lab techniques. I would also like to thank Wilfrid Laurier University and the Biology department as well as the grad students within the department who all made this experience a very positive one.

## Table of Contents

ABSTRACT .....	ii
ACKNOWLEDGEMENTS .....	iv
TABLE OF CONTENTS .....	v
LIST OF FIGURES .....	vii
LIST OF ABBREVIATIONS .....	viii
1. INTRODUCTION .....	1
1.1 Plastids .....	1
1.2 Protein targeting .....	4
1.3 Protein import into chloroplasts .....	6
1.4 Chloroplast translocation machinery .....	7
1.4.1 The Tic complex .....	7
1.4.2 The Toc complex .....	11
1.5 The Toc GTPases in <i>Arabidopsis thaliana</i> .....	15
1.6 Biophysical approaches to studying protein-protein interactions .....	18
1.6.1 Fluorescence spectroscopy .....	18
1.6.2 Fluorescence spectroscopy and proteins .....	19
1.6.3 Circular dichroism spectroscopy .....	22
1.7 Objectives and hypotheses .....	26
1.7.1 Observe interaction events between atToc159G and atToc33G using biophysical techniques .....	26
1.7.2 Manipulation of the monomer/dimer relationship between atToc159G and atToc33G .....	27
2. METHODS .....	28
2.1 Construction of wild-type and mutant atToc159G and atToc33G constructs .....	28
2.1.1 Wild-type atToc33G and atToc159G constructs .....	28

2.1.2 Mutant atToc159G W973F and W1056F mutant constructs .....	28
2.2 Expression and purification of recombinant wild-type atToc33G <sub>His</sub> and wild-type and mutant atToc159G <sub>His</sub> proteins .....	30
2.2.1 Wild-type atToc159G <sub>His</sub> and atToc33G <sub>His</sub> .....	30
2.2.2 AtToc159G single tryptophan mutant proteins .....	32
2.3 Nucleotide removal by chelating Mg <sup>2+</sup> cofactor .....	32
2.4 Blue Native PAGE .....	33
2.5 Fluorescence Spectroscopy.....	33
2.6 Circular Dichroism Spectroscopy .....	33
3. RESULTS .....	35
3.1 Homology model of atToc159G .....	35
3.2 Expression and purification of atToc159G, atToc33G, and atToc159G single tryptophan mutants .....	35
3.3 Structure of wild-type atToc159G and atToc33G and their interaction .....	39
3.4 Structure of atToc159G single tryptophan mutants .....	44
3.5 EDTA and Chelex treatment of wild-type proteins atToc159G and atToc33G .....	47
4. DISCUSSION .....	55
4.1 Homodimerization of atToc33G and atToc159G .....	55
4.2 Heterodimer formation is an uncommon event <i>in vitro</i> .....	57
4.3 AtToc159G tryptophan mutations .....	59
4.4 Future investigations .....	60
4.5 Integrating chloroplast protein translocation into the biological sciences .....	61
5. CONCLUSIONS .....	63
REFERENCES .....	64
APPENDIX 1 Primers for generating single tryptophan mutants of atToc159G using PCR .....	72
APPENDIX 2 Calculation of monomer:dimer ratio for BN-PAGE using ImageJ .....	73

## List of Figures

**Figure 1.** Different types of plastids.

**Figure 2.** Structure of chloroplast.

**Figure 3.** Tic Toc import pathway.

**Figure 4.** Schematic diagram of the Tic and Toc complexes.

**Figure 5.** Schematic diagram of Toc159 and Toc34 families of proteins from *Arabidopsis thaliana*.

**Figure 6.** Transition states showing absorption and fluorescence.

**Figure 7.** Typical fluorescence absorption and emission spectra for aromatic amino acids.

**Figure 8.** Effect of tryptophan environment on emission spectra.

**Figure 9.** Far UV spectra showing different secondary structures.

**Figure 10.** Amino acid sequence of atToc159G.

**Figure 11.** Crystal structure of atToc33G and homology model of atToc159G.

**Figure 12.** SDS-PAGE of expression and purification of recombinant proteins atToc159G and atToc33G.

**Figure 13.** SDS-PAGE of atToc159G single tryptophan mutants.

**Figure 14.** CD and Fluorescence spectra of atToc159G (5  $\mu$ M) and atToc33G (5  $\mu$ M) at 20°C and pH 7.5.

**Figure 15.** Blue native PAGE of wild-type atToc33G and wild type and mutant atToc159G.

**Figure 16.** CD and fluorescence spectra showing atToc159G and atToc33G interaction.

**Figure 17.** Analysis of the interaction between atToc159G and atToc33G using BN-PAGE.

**Figure 18.** CD and Fluorescence spectra of atToc159G single tryptophan mutants (5  $\mu$ M) at 20°C and pH 7.5.

**Figure 19.** EDTA treatment of 5  $\mu$ M atToc159G and atToc33G.

**Figure 20.** CD and FL spectra of 2.5  $\mu$ M wild-type and chelex treated atToc159G and atToc33G.

**Figure 21.** BN-PAGE of chelex treated atToc159G and atToc33G.

**Figure 22.** Effects of Mg<sup>2+</sup> and nucleotide on chelex treated atToc33G.

**Figure 23.** Effects of Mg<sup>2+</sup> and nucleotide on chelex treated atToc159G.

## List of Abbreviations

ATP: adenosine tri-phosphate

atToc159G W1056F: atToc159G single tryptophan mutant where tryptophan 1056 is mutated to phenylalanine

atToc159G W973F: atToc159G single tryptophan mutant where tryptophan 1056 is mutated to phenylalanine

BFA: Brefeldin A

BiFC: bi-molecular fluorescence complementation

BN-PAGE: blue native poly-acrylamide gel electrophoresis

CD: circular dichroism

ceQORH: quinone oxidoreductase

CuCl<sub>2</sub>: Copper Chloride

EDTA: ethylenediaminetetraacetic acid

ER: endoplasmic reticulum

FL: fluorescence

FRET: fluorescence resonance energy transfer

GTP: guanosine tri-phosphate

Hsp90/70: heat shock protein 90 or 70

IEM: inner envelope membrane

IMAC: immobilized metal affinity chromatography

kDa: kilodaltons

NPP: nucleotide pyrophosphatase/phosphodiesterase

OEM: outer envelope membrane

OEP64: outer envelope protein 64

PEG: polyethylene glycol

SPP: stromal processing peptidases

SRP: signal recognition particle

THP: Tris(hydroxypropyl)phosphine

Tic: translocon at the inner envelope membrane of chloroplasts

Toc: translocon at the outer envelope membrane of chloroplasts

Tom40: translocation at the outer envelope membrane of the mitochondria protein 40

TPR: tetratricopeptide repeat

UV: ultraviolet

## **1. Introduction**

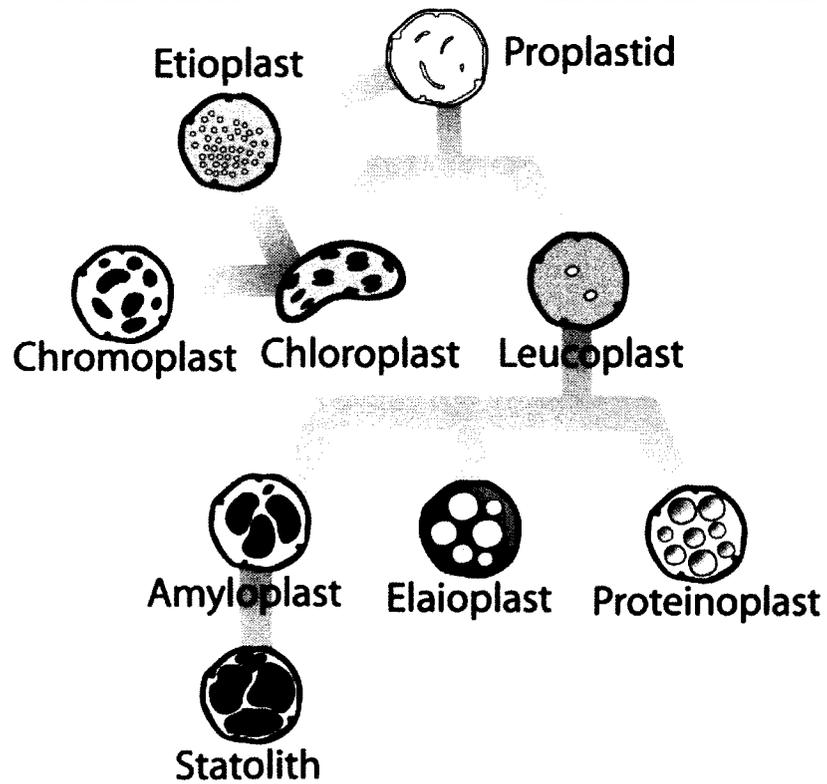
### **1.1 Plastids**

The evolutionary lineage of cellular compartments known as plastids has given rise to a family of organelles necessary for the proliferation of plant life. Plastids are responsible for performing a diverse number of tasks in plant tissues including the storage of cellular materials, amino acid and lipid production as well as other important processes such as photosynthesis (Wise and Hooper, 2007). Proplastids, the progenitor organelle to all other plastid types, give rise to specific plastids such as amyloplasts, which store starch; elaioplasts, which store fats; chromoplasts, which synthesize and store pigments; statoliths, responsible for detecting gravity in roots; and chloroplasts, which are the green pigmented photosynthetic organelle (Figure 1; Wise, 2006). Each plastid type is partly defined by the protein complement located within the organelle, which maintains the ability to interconvert between types depending on the proteins present (Roberts, 2007). This dynamic structure is directly linked to the functional roles that the protein complement plays in the metabolic pathways characteristic of each plastid type.

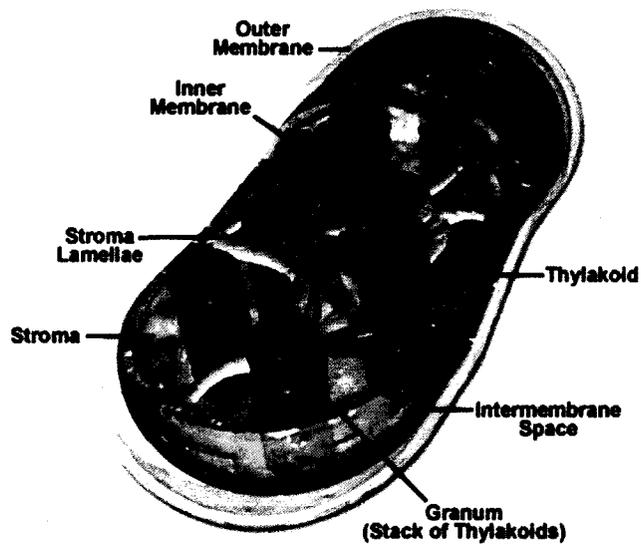
Plastids are thought to have originated from an endosymbiotic event between a eukaryotic-like cell and a cyanobacterium around 1.2 to 1.5 billion years ago (Dyall et al., 2004). The incomplete digestion of the cyanobacterium by the host cell resulted in the formation of an organelle surrounded by a double membrane, which is characteristic of all plastids today. Despite this common double membrane structure among plastids, the internal composition varies greatly. For example, chloroplasts maintain a complex inner structure consisting of inner membranes further compartmentalizing the organelle. More specifically, chloroplast structure is made up of an outer and inner membrane surrounding a stromal space which contains the thylakoid membrane and a lumen space within (Figure 2). Amyloplasts, on the other hand, contain less inner membranous structure, but are mostly filled with starch granules (Figure 1; Wise, 2006). Following the endosymbiotic event a reduction in genome size as a result of gene transfer from

## Plastids

---



**Figure 1.** Different types of plastids. Proplastids are the progenitor for all other plastid types. Other plastid types are shown with relationships between them in blue lines. Figure taken from Wikiversity online cell biology image database.



**Figure 2.** Structure of chloroplast. 3D model of chloroplast showing the outer and inner membrane as well as the internal thylakoid membranes and stromal lamellae (Image taken with permission from <http://micro.magnet.fsu.edu>).

the endosymbiont to the host nucleus occurred, such that 90-95% of all plastid proteins are now encoded in the nucleus (Dyall et al., 2004; Martin et al., 1998). Experimentally, the incidence of gene transfer has been shown to have occurred multiple times in over 300 angiosperms with respect to translation initiation factor *infA*, which has been noted to be the most mobile chloroplast gene in plants (Millen et al., 2001). It has also been argued that plastid development may be a result of convergent evolution and not divergence as previously thought (for a review of both theories see Palmer, 2003; Stiller et al., 2003). Overall, plastid development shows the adaptive nature of plants including the development of unique organelles that have become an integral part of many diverse plant species.

## 1.2 Protein targeting

The evolutionary transition from free-living cyanobacterium to a semi-autonomous, integrated organelle, such as the chloroplast, necessitated the evolution of a protein trafficking system to enable the import of proteins needed for organelle biogenesis and maintenance from the cytosol. Since the specific and efficient targeting and import of such compartment-destined proteins is required for life, it has naturally become a noteworthy field of research. Protein targeting has been well studied and characterized in many cellular systems including the endoplasmic reticulum (ER), mitochondria, and chloroplasts. Early research led to the development of the signal hypothesis, which now acts as the basis for protein import mechanisms (Dobberstein and Blobel, 1975a/1975b). The signal hypothesis states that targeting information is contained within the protein to be targeted and not in chaperones and translocons. Investigating the signal recognition particle (SRP) pathway, Dobberstein and Blobel (1975 a/1975b) showed that unique signals encoded within the mRNA, which are translated to the nascent chain of the protein are recognized by the SRP and promote binding of the translating ribosome to the ER, thus delivering nascent polypeptides with a signal sequence to the ER. It was also shown that the N-terminal signal sequence was subsequently cleaved generating mature protein without the signal sequence. Such signal sequences (also referred to as transit peptides, or presequences, depending on the translocation pathway they are part of) are responsible for the targeting of proteins to specific locations within cells.

Transit peptides direct targeting of many proteins to their appropriate location but also promote translocation and targeting by helping to maintain a translocation appropriate unfolded or partially folded state of the protein (Bruce, 2000). In general, folded proteins do not have the capacity to pass through mitochondrial and chloroplast translocation channels, as they are too large. It has been observed that the pores formed by the mitochondrial channel protein Tom40 and chloroplast channel protein Toc75 are 20 Å and 14 Å in diameter, respectively, whereas proteins average from 20 – 60 Å in diameter when folded (Künkele et al., 1998; Hinnah et al., 2002; Sundararaj et al., 2004). However, studies have shown that Toc75 has the ability to translocate a 23 Å folded fusion protein, prOE17-BPTI, revealing its dynamic capacity (Clark and Theg, 1997). Preproteins are maintained in an unfolded or linear conformation, presumably by the action of chaperones, making it possible for proteins to pass through the translocation pores of chloroplasts and mitochondria (Bruce, 2000). Another role of some transit peptides is to interact with chaperones increasing the rate of targeting. Outer envelope protein 64 (OEP64) has been shown to interact with heat shock protein 90 (Hsp90), which is thought to be necessary for the targeting of transit peptides to the chloroplast translocation complex in some cases (Qbadou et al., 2006). It has been proposed that OEP64 recognises chaperone-associated preproteins and mediates the introduction of the preprotein into the translocation machinery of the chloroplast (Qbadou et al., 2006). However, *in vivo* data has shown that OEP64 is not essential for the import of proteins into chloroplasts in *Arabidopsis thaliana* (Aronsson et al., 2007). Plants containing knockouts of all three OEP64 genes (III, V, and I) showed no distinguishable differences from wild-type with respect to chlorophyll accumulation, photosynthetic performance, organellar ultrastructure and chloroplast protein accumulation (Aronsson et al., 2007). Therefore, the importance of OEP64 for chloroplast protein import remains in question. Despite current controversies surrounding the series of events that promote protein targeting, transit peptides are the foundation of sub- and extra-cellular compartmentalization of proteins. Successful targeting of proteins is only a part of the journey as recognition by the translocation machinery as well as the translocation event itself are next in the import pathway.

### 1.3 Protein import into chloroplasts

Protein import is a complex and multistep process, which has been well studied in mitochondria (for a review see Baker et al., 2007) and more recently in chloroplasts (for a review see Smith, 2006). Both involve multi-protein interactions and specific energy requirements of adenosine triphosphate (ATP) for mitochondria and ATP and guanosine triphosphate (GTP) for chloroplasts (Mokranjac and Neupert, 2008; Smith, 2006). For the purposes of this thesis, focus will remain on the protein import machinery of the chloroplast. Most notably the site of photosynthesis, chloroplasts contain a genome that is approximately 150 kilobases with approximately 87 protein encoding regions (Sato et al., 1999). According to the endosymbiotic hypothesis, the small size of the chloroplast genome is a result of gene transfer from the chloroplast to the nucleus (Dyall et al., 2004). This transfer necessitated the evolution of specific import machinery for chloroplast destined proteins. The Tic and Toc complexes have been shown to import both photosynthetic and non-photosynthetic proteins and make up the major import pathway for chloroplasts (Smith, 2006).

Recent discoveries have revealed possible alternative import pathways for chloroplast-destined proteins. One such pathway uses a currently uncharacterized protein translocation system. Import of chloroplast inner envelope protein quinone oxidoreductase (ceQORH) has been shown to be imported using a proteinaceous receptor, however, import is unaffected by competition with the Toc-Tic substrate precursor of ferredoxin or Toc component antibody, revealing an alternative import mechanism other than the Toc complex (Miras et al., 2007). Also, the import of Tic32 is not regulated by Toc159, Toc34, or Toc75-III and import is not deterred by treatment with thermolysin or treatment with  $\text{CuCl}_2$  or spermine, which are known Toc75 inhibitors (Nada and Soll, 2005). Evidence for protein import into chloroplasts via the secretory pathway has also been shown (Nanjo et al., 2006; Villarejo et al., 2005). Chloroplast protein nucleotide pyrophosphatase/phosphodiesterase (NPP) has been shown to be localised to the chloroplast using green fluorescent protein as well as being N-glycosylated using Endo-H digests. Since

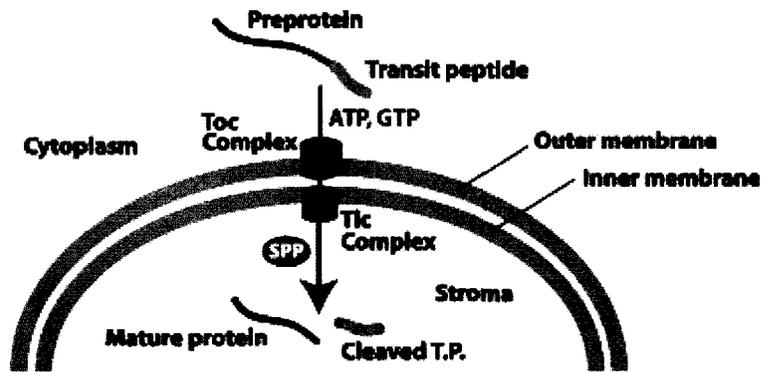
glycosylation occurs solely in the ER, Endo-H digestion identifies proteins that have been glycosylated by specifically cleaving asparagine-linked mannose-rich oligosaccharides. The amount of glycosylation can also be quantified by subsequent protein analysis. Also, treatment with Brefeldin A (BFA), a metabolite of fungus which inhibits protein transport from the ER to the Golgi, results in localisation of NPP to the ER suggesting a vesicular transport of NPP from the ER to the chloroplasts (Nanjo et al., 2006). Although the Tic/Toc import system may be the best characterised pathway for the import of proteins into the chloroplasts it is clear that there are alternative pathways which act through both the secretory pathway and other unknown proteinaceous import pathways.

In terms of the Tic/Toc complex, nuclear encoded chloroplast-destined proteins are translated in the cytoplasm and contain an N-terminal transit peptide which targets them to the Toc complex. This transit peptide is recognized by the Toc complex which promotes translocation in a unidirectional fashion. In coordination with the Toc complex, the Tic complex completes translocation across the double membrane and the transit peptide is cleaved by a stromal processing peptidase either revealing a secondary transit peptide for targeting the protein within the chloroplast, or allowing for folding to occur, thereby allowing the protein to become functional (Figure 3; Smith and Schnell, 2004; Smith, 2006).

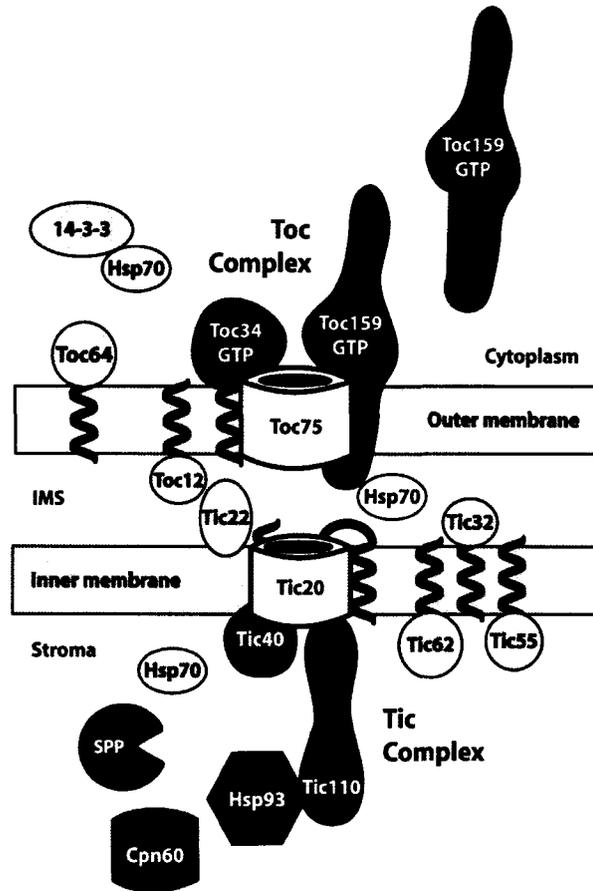
## **1.4 Chloroplast translocation machinery**

### **1.4.1 The Tic complex**

A schematic diagram of the Tic/Toc complex as was originally characterised based on biochemical approaches in *Pisum sativum* is shown in Figure 4. Each protein is named for its location with either the Tic or Toc complex as well as its size in kilodaltons (kDa) (Schnell et al., 1997). Translocation across the inner envelope membrane (IEM) is less well characterised than the outer envelope membrane (OEM) but certain details have been established. The major components of the Tic



**Figure 3.** Tic/Toc import pathway. Chloroplast destined proteins contain an N-terminal transit peptide targeting them to the Toc complex which promotes translocation of the preprotein through to the stroma via the Tic complex where stromal processing peptidases (SPP) cleaves the transit peptide to allow for proper folding of the mature protein. Figure taken from Smith and Schnell, 2004.



**Figure 4.** Schematic diagram of the Tic and Toc complexes. Figure taken from Smith, 2006.

complex are Tic110, Tic40, Tic20, and Tic22, with associated proteins Tic62, Tic32, and Tic55 (Figure 4). The inner pore complex which forms the translocation channel has been hypothesised to be formed by a variety of proteins and remains a controversial topic in the literature. Evidence has been presented showing that Tic110 forms a  $\beta$ -barrel pore channel approximately 15 Å wide when reconstituted into liposomes (Heins et al., 2002). However this data has been disputed; evidence has been presented that supports a soluble  $\alpha$ -helical C-terminus of Tic110, which was in agreement with findings on this protein from prior to 2002 (Inaba et al., 2003, and references therein). Data showing that Tic110 is necessary for plant development reveals the critical role of Tic110 in the import pathway (Kovacheva et al., 2005); however, evidence for the role of Tic20 in channel formation also exists. Tic20 has been shown to contain four transmembrane  $\alpha$ -helices when originally characterised in pea and is critical for plant development (Kouranov et al., 1998). Using antisense lines in *Arabidopsis*, Tic20-deficient plants were pale-yellow and showed reduced translocation at the inner envelope membrane (Chen et al., 2002). Similar to Tic20 topology, and most recently discovered, Tic21 has been shown to be critical for import, and *Arabidopsis thaliana* Tic21 null mutants accumulate preproteins outside the chloroplast (Teng et al., 2006). Other studies have challenged the idea that Tic21 is involved in protein translocation and suggest that it may act as an iron transporter and help regulate cellular metal homeostasis (Duy et al., 2007). Recent evidence has been presented that supports the role of Tic20 and Tic21 as likely components of the inner membrane channel (Kikuchi et al., 2009). It is clear that much more work needs to be done in order to adequately model the translocation of proteins through the inner envelope membrane.

Associating proteins like Tic40, Tic62, Tic55, and Tic32 are not necessary for plant development; however, they are thought to assist in the efficiency and regulation of translocation at the IEM. Tic40 is a protein consisting of a single transmembrane domain at the N-terminus and a soluble C-terminus that extends into the stroma (Chou et al., 2003). It has been shown to cross-link to Tic110 and also contains a tetratricopeptide repeat (TPR) domain; such domains are often involved in protein-protein interactions (Stahl et al., 1999). Tic40 acts as a co-chaperone with Hsp93 to hydrolyse ATP and this is thought to

promote translocation in an ATP-dependent manner (Chou et al., 2003; Chou et al., 2006), thereby potentially explaining the ATP requirement for chloroplast protein import. Tic62, Tic55, and Tic32 are redox proteins that have been found to associate with the Tic complex in some instances (Caliebe et al., 1997; Hormann et al., 2004; Stengel et al., 2008). Motifs characteristic of NAD(P)H dehydrogenases can be found at the N-terminus of Tic62 while the C-terminus has been shown to interact with ferredoxin-NAD(P)<sup>+</sup> oxidoreductase indicating its potential role in regulating translocation by sensing redox state (Küchler et al., 2002). Tic55 is thought to bind iron by a Rieske-type iron-sulphur centre characteristic of electron transfer, thereby contributing to the regulation of the Tic complex based on the redox state (Caliebe et al., 1997). Tic32 is a calmodulin-binding protein suggesting a significant role for calcium in the regulation of protein import (Chigri et al., 2006). It is noteworthy, however, that none of Tic62, Tic55 or Tic32 have been shown to play a direct role in import; therefore, whether they are bona fide “Tic” proteins remains somewhat in question.

#### **1.4.2 The Toc complex**

Figure 3 shows the three major components of the Toc complex (Toc159, Toc34, and Toc75). Toc75 is the channel protein and Toc34 and Toc159 are GTPases that act as preprotein receptors which also facilitate translocation (Smith, 2006). Recent studies using Blue Native polyacrylamide gel electrophoresis (BN-PAGE) have been used to support a stoichiometry of 1:3:3 for Toc159:Toc75:Toc34, however, other studies have suggested stoichiometries of 1:5:4 and 1:5:2 using other experimental techniques (Kikuchi et al., 2006, Schleiff et al., 2003; & Vojta et al., 2004). Therefore, the structural specifics of the Toc complex remain unclear. However, the common thread does indicate a single Toc159 protein acting in concert with multiple Toc75 channels and multiple Toc33 receptors, revealing the dynamic nature of Toc159 during the translocation process.

## **Toc75**

Structurally, Toc75 is a  $\beta$ -barrel protein consisting of 18 amphiphilic  $\beta$ -strands which creates a pore approximately 14 Å in diameter (Schleiff et al., 2003; Hinnah et al., 2002). Homologues of Toc75 have been identified in gram negative cyanobacterium *Synechocystis* sp. PCC 6803 showing similar structure (Reumann et al., 2005). The Toc75 channel is related to an ancient outer envelope membrane protein family found in gram negative bacteria, outer membrane protein 85 (OMP85; Reumann et al., 2005). As gene transfer occurred post-endosymbiosis, topical inversion of the OMP85 export channel may have resulted in the Toc75 import channel seen today (Reumann et al., 2005). Three potential mechanisms for the gating of Toc75 have also been hypothesised (Li et al., 2007). Li et al. (2007) hypothesise that Toc75 may be tightly linked to the inner membrane translocation channel in an open constitutive open state which is then gated by Toc34 and Toc159 depending on conformational changes of the protein resulting in changes to the dimer formation or potentially the hinge region between G-domains and the membrane anchor. The second gating mechanism proposes that localized membrane potential may allow Toc75 to be voltage-gated. Studies have shown that Toc75 closes when membrane potentials of  $\pm 75$  mV occur (Hinnah et al., 2002). An unknown mechanism may be responsible for the change in membrane potential allowing Toc75 to open and close as needed. The third mechanism proposes that Toc75 remains in an open state all the time and is not shielded by Toc34 or Toc159. The inner membrane channel of the Tic complex would then be the primary gate and may be stimulated by the introduction of transit peptides into the intermembrane space or by some other translocon component (Li et al., 2007).

## **Toc34**

Toc34 is a GTPase receptor, which contains a single transmembrane domain anchoring the protein to the chloroplast outer membrane while the remainder of the protein occurs in the cytosol (Kessler et al., 1994). This cytosolic domain contains the GTP binding site as well as a putative protein-

protein interaction site (Tsai et al., 1999; Koenig et al., 2008). The insertion of Toc34 into the outer chloroplast membrane is mediated by both a proteinaceous and an ATP-dependent component (Tsai et al., 1999). GTP requirements for insertion of Toc34 have also been investigated and show that GTP is only necessary for maintaining a membrane insertion efficient folded state of full-length protein (Chen and Schnell, 1997; Tsai et al., 1999). Toc34 also contains a targeting sequence necessary for insertion which has been shown to allow for the translocation of a 26 kDa recombinant protein containing the targeting sequence across the outer membrane (Li and Chen, 1997). It has also been shown that this targeting sequence is located partially within the transmembrane domain as well as the GTP binding domain (Chen and Schnell, 1997).

### **Toc159**

Toc159 is anchored to the chloroplast outer envelope membrane, but interestingly, no transmembrane domains are predicted within its amino acid sequence (Kessler et al., 1994); how it is anchored in the membrane, therefore, is unknown. The remainder of the protein contains both an acidic domain shown to be intrinsically unstructured (Richardson, 2008) and a GTPase domain (Bauer et al., 2000). The insertion of atToc159 into the outer membrane of chloroplasts has also been shown to be mediated by its G-domain and to require GTP (Smith et al., 2002). Toc159 has been shown to interact directly with Toc34 proteins in both pea and *Arabidopsis thaliana*, which promotes its insertion into the outer membrane (Hiltbrunner et al., 2001; Wallas et al., 2003). Toc159 and its occurrence in the Toc complex as indicated by its direct interaction with Toc34 reveals its important role in the translocation of proteins across the outer envelope membrane of the chloroplast.

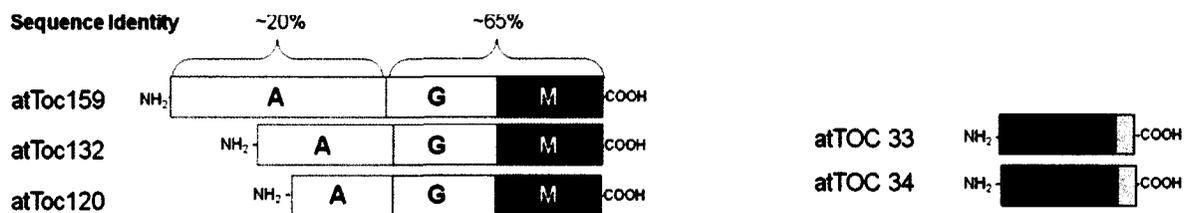
Currently two major hypotheses, the ‘targeting’ hypothesis and the ‘motor’ hypothesis, have been developed to explain the mechanism of preprotein targeting to the Toc complex. The ‘targeting’ hypothesis proposes that Toc159 serves as the primary preprotein receptor and GTP-bound Toc159 and Toc34 act co-operatively to regulate translocation (reviewed in Kessler & Schnell, 2006). Hydrolysis of

GTP to GDP is activated by preprotein binding and confers a conformational change in the Toc components promoting insertion of the preprotein into the translocation channel. Once translocation has been started, unidirectional import is maintained by an ATP-dependent cycle between the preprotein and intermembrane space Hsp70 (Kessler & Schnell, 2006). The alternative hypothesis, the ‘motor’ hypothesis, proposes that Toc34 bound with GTP is the primary preprotein receptor and binds a phosphorylated form of the preprotein transit peptide (Becker et al., 2004). Binding of the preprotein to Toc34 induces GTP hydrolysis and the preprotein is transferred to Toc159-GTP. Dephosphorylation of the preprotein occurs and a GTP-GDP cycling of Toc159 acts to propel the protein through the translocation channel (reviewed in Kessler & Schnell, 2006). Two other translocation models have been proposed that consider more recent data. The occurrence of homodimeric structures for both Toc159 and Toc34 in pea has been used to generate a model which depicts the use of homodimers for protein translocation (Li et al., 2007). The model hypothesizes that GDP-bound Toc159 occurs as a homodimer and is capable of binding precursors. Precursor binding induces the exchange of GDP to GTP which in turn induces the GDP to GTP exchange of Toc34-GDP homodimer. This exchange may be a result of Toc159-Toc34 or even Toc159-Toc75-Toc34 interactions. Toc159-GTP then interacts with Toc34-GTP which induces GTP hydrolysis and this energy is used to promote translocation to the trans side of Toc75 (Li et al., 2007). Most recently, another model, which acts to reconcile the ‘motor’ and ‘targeting’ hypotheses has been proposed (Agne and Kessler, 2009). This consensus model proposes that chaperones such as Hsp70/14-3-3 and Hsp90 guide preproteins to Toc34 and Toc159, which are located in the Toc complex. GTP-bound Toc34 and Toc159 act cooperatively as a gating mechanism, which promotes the insertion of the preprotein into the Toc75 channel upon GTP hydrolysis. Subsequently, transmembrane space Hsp70 may contribute to translocation across the outer membrane. Recently discovered Toc12 may then recruit Hsp70 from within the intermembrane space via its J-motif to also promote translocation. Toc34 and Toc159 would then reset themselves to their GTP loaded states in preparation of further translocation cycles (Agne et al., 2009; Agne and Kessler, 2009).

## 1.5 The Toc GTPases in *Arabidopsis thaliana*

In *Arabidopsis thaliana*, multiple homologues have been identified for Toc159 and Toc34. The Toc159 gene family is made up of four homologues (atToc159, atToc132, atToc120, and atToc90) and the Toc34 gene family is made up of two (atToc34 and atToc33) (Jarvis et al., 1998; Bauer et al., 2000). Figure 5 is a schematic diagram of the Toc159 and Toc34 gene families in *Arabidopsis thaliana*. This diagram shows the tripartite structure of the Toc159 gene family. Each protein in this family contains an N-terminal acidic (A) domain, a C-terminal membrane (M) bound domain and an internal GTPase (G) domain (Bauer et al., 2000). AtToc33 and atToc34 both contain C-terminal  $\alpha$ -helical membrane anchors but are mostly dedicated GTPases.

Recent studies have provided evidence to suggest that structurally and functionally distinct Toc complexes exist (Ivanova, et al., 2004; Kubis et al., 2004). Ivanova et al. (2004) showed that Toc complexes containing atToc159 also contain atToc33 and ones containing atToc132/atToc120 contain atToc34. AtToc159 containing complexes have also been shown to be responsible for the import of photosynthetic proteins, whereas atToc132/atToc120 containing complexes import non-photosynthetic proteins (Ivanova et al., 2004; Smith et al., 2004; Lee et al., 2009). Research has also shown that both homotypic and heterotypic dimers form in vitro (Hiltbrunner et al., 2001; Sun et al. 2002; Ivanova et al., 2004; Yeh et al., 2007). These dimers occur via the G-domains of both Toc159 and Toc34 gene families suggesting that distinct Toc complexes may be formed through interaction between these domains (Ivanova et al., 2004). Despite the evidence to suggest the formation of different complexes, confirmation of such phenomenon using biophysical techniques to assess protein-protein interactions between these proteins has not been attempted.



**Figure 5.** Schematic diagram of Toc159 and Toc34 families of proteins from *Arabidopsis thaliana*. Toc159 family of proteins show a tripartite structure with acidic (A), GTP binding (G) and membrane bound (M) domains. Sequence identity for Toc159 proteins is given above the A and GM domains (Ivanova et al., 2004). Toc34 family of proteins shows the GTP binding domain and small transmembrane domain.

The crystallization of *Pisum sativum* Toc34<sub>GDP</sub> revealed a dimeric structure thought to involve an arginine finger supplied by the Toc34's dimeric partner (Sun et al., 2002). However, more recent evidence has been presented to show that mutations in this arginine finger do not inhibit GTP hydrolysis using *in vitro* GTP hydrolysis assays, suggesting that the arginine finger may only contribute to the formation of homodimers (Weibel et al., 2003; Agne and Kessler, 2009; Lee et al., 2009). Recent crystal structures for psToc34<sub>GMPPNP</sub>, atToc33<sub>GDP</sub>, and atToc33<sub>GMPPNP</sub> (PDB 3BB1, 3BB3, 3BB4, respectively) have been solved and hypotheses proposing switch I catalytic activity for GTP hydrolysis are discussed (Koenig et al., 2008). Regardless of mechanism, the importance of protein-protein interactions between members of the Toc34 gene family has been firmly established (Koenig et al., 2008; Lee et al., 2009). Theories involving possible homodimerization followed by subsequent heterodimer formation between members of the Toc34 and Toc159 gene families have been proposed, but the significance of such interactions have yet to be validated. Possible transit peptide binding sites in psToc34 have also been identified by Koenig et al. (2008). A 63 Å<sup>3</sup> cavity located between helix α1, helix α0, and the central β sheet crystallized with a molecule of polyethylene glycol (PEG) in it in some cases. This artifact of the crystallization process is thought to be an indication of possible substrate binding or protein-protein interaction site (Koenig et al., 2008). Studies focussed on elucidating the importance of this cavity would be of great benefit to our understanding of protein import into chloroplasts.

Observations of heterodimeric interactions between the G-domains of Toc159 and Toc34 proteins have recently been made. Both yeast two-hybrid and split ubiquitin systems have been used *in vitro* and *in vivo* to show the interaction between atToc159G and atToc33G (Rahim et al., 2008). Interestingly, detection of homodimers was not possible using the yeast two-hybrid system which suggests that the heterodimerization between Toc159 and Toc34 proteins is important. Importantly, the observation of heterodimerization *in vivo* confirms many previous models regarding the importance of Toc159-Toc34 interactions for successful protein translocation (Rahim et al., 2008). The development of the plant cell-

based split ubiquitin system for studying protein-protein interaction has been shown to have great utility and future studies should focus on its further development.

## **1.6 Biophysical approaches to studying protein-protein interactions**

### **1.6.1 Fluorescence spectroscopy**

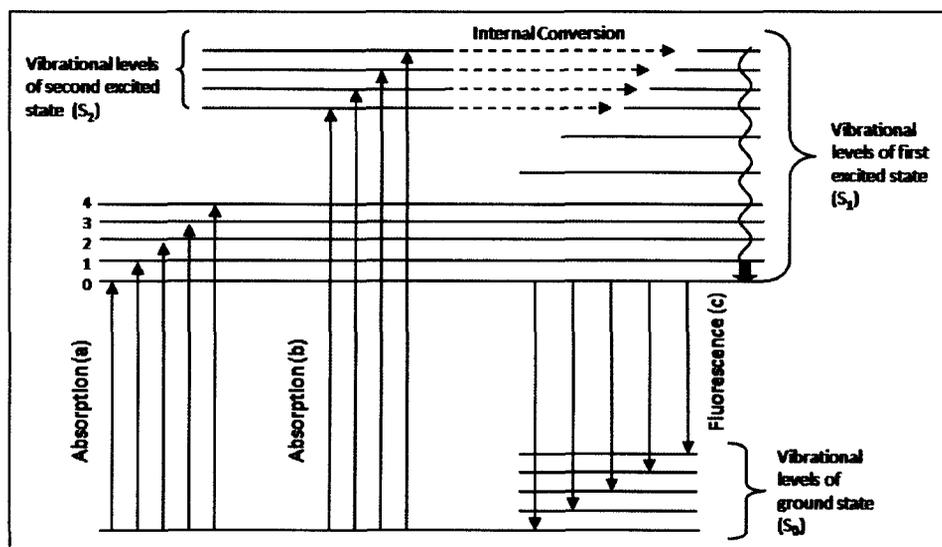
The development of techniques using fluorescent properties of both naturally occurring fluorescent molecules as well as developed fluorophores has widespread use in research today. The relatively low sample concentrations that are used, high sensitivity and dynamic range of fluorescent biomolecules have made the technique of fluorescence spectroscopy useful and efficient (Szabo, 2000). Specifically, techniques such as steady-state fluorescence spectroscopy, fluorescence resonance energy transfer (FRET) and bi-molecular fluorescence complementation (BiFC) have been developed that can provide valuable information about proteins and their surrounding environment (Sharma and Kalonia, 2003; Piston and Kremers, 2007; Sung and Huh, 2007). The use of extrinsic fluorophores, such as fluorescein, for the study of biomolecules can often be advantageous as it allows for increased signal detection sensitivity. However, such probes are introduced randomly, and their size and chemical properties can often affect the structure of the biomolecule being studied (Ross et al., 1997). More sophisticated methods have been used to introduce fluorophores and include the development of complex tRNA substitutes for specific codons as well as the creation of coding sequences which introduce such fluorophores into specific sites (Twine and Szabo, 2003). One such example is 5-hydroxytryptophan, which enables specific detection of proteins (Twine and Szabo, 2003). Alternatively, naturally occurring or intrinsic biomolecules, such as tryptophan, can be used to gather similar information without the high cost and extraneous sample preparation.

Briefly, and in general, fluorescence spectrometers utilize as a light source ultra-violet and visible light (UV;  $\lambda = 200\text{nm} - 800\text{nm}$ ), which is passed through a solution containing known samples; fluorescence that is emitted from the sample can be detected and measured (Lakowicz, 2006). For all

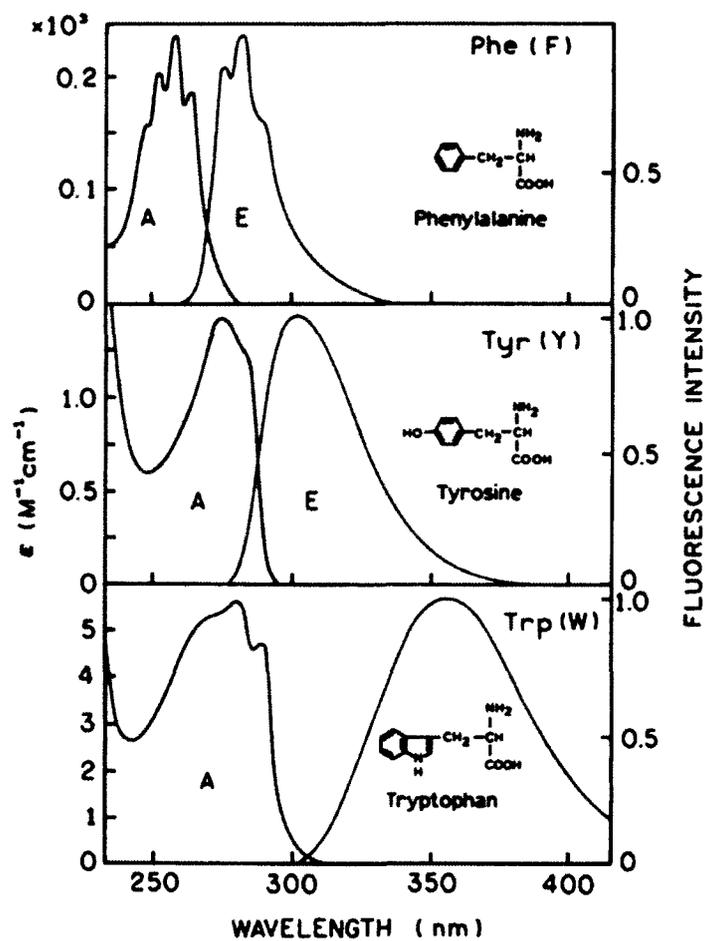
fluorescent molecules, light is absorbed by the solution producing excited electron states. This excitation causes the electrons to establish excited transitional states (Fig. 6, absorption (a) and (b)). Figure 6 shows two possible excited states,  $S_1$  and  $S_2$ , during absorption events (a) and (b), respectively. Within these transitional states exists a series of vibrational levels, which correspond to different electron energy levels within the same state (Parker, 1968). Often internal conversion can be observed as molecules shift from low vibrational levels of a higher energy state to higher vibrational levels of a lower energy state, aiding in the process of re-stabilization. The energy released during this time can be accounted for by a number of pathways. These pathways include Stokes' shift, non-radiative decay, intersystem crossing, and radiative or fluorescent decay (Parker, 1968). As shown in Figure 6, fluorescence (c) is emitted during the relaxation of molecules from an excited state back down to the ground state.

### **1.6.2 Fluorescence spectroscopy and proteins**

In more complex systems, like that of proteins, fluorescence spectroscopy has been used in structure-function studies. Intrinsic fluorescence of proteins is often an effective way of studying protein structure and environment and is made possible by the aromatic amino acids tyrosine and tryptophan, which absorb light in the UV spectral range. Emission spectra typical of all the aromatic amino acid is shown in figure 7. More specifically, tryptophan fluorescence has been used extensively in studies partially due to its significant extinction coefficient,  $\epsilon_{280} = 5579 \text{ cm}^{-1}$  (Szabo, 2000). Tryptophan is also extremely sensitive to its microenvironment making it a suitable candidate for detecting small changes to protein structure (Lakowicz, 2006). Early experimental work showed that tryptophan environments within proteins can be estimated based on the emitted fluorescence maximum and spectral band width and could predict the fraction of fluorescence that is quenched by ionic quenchers, quantum yields, and fluorescence lifetimes (Burstein et al., 1973). Since then, characteristic fluorescence emission spectra for tryptophan residues within a protein's 3D structure have been developed based on protein conformation and relative tryptophan position (Figure 8; Lakowicz, 2006).



**Figure 6.** Transition states showing absorption and fluorescence. Two absorption scenarios (a) and (b) are shown both resulting in fluorescence emission (c). The vibrational levels of both the ground state ( $S_0$ ) and excited states ( $S_1$  and  $S_2$ ) are shown. Arrows indicate the change in electronic configuration where energy is on the y-axis. Modified from Parker, 1968.

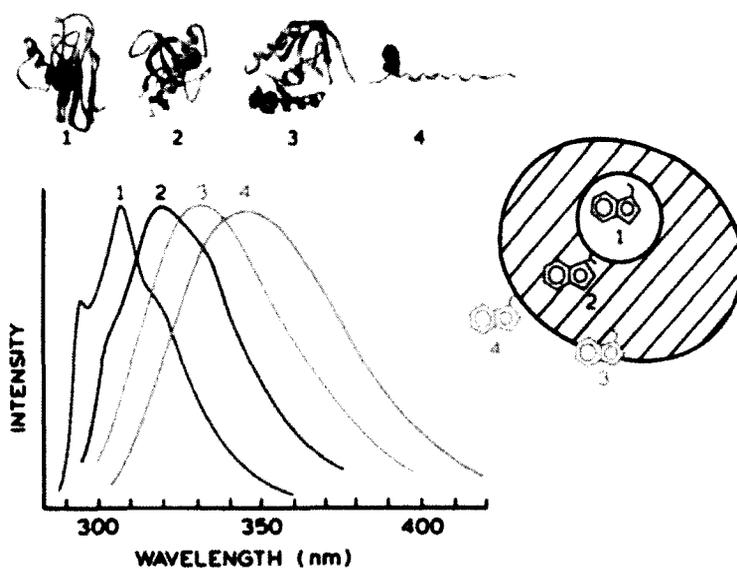


**Figure 7.** Typical fluorescence absorption and emission spectra for aromatic amino acids. Absorption (A) and emission spectra (E) for phenylalanine, tyrosine, and tryptophan at pH 7.0. Figure taken from Lakowicz, 2006.

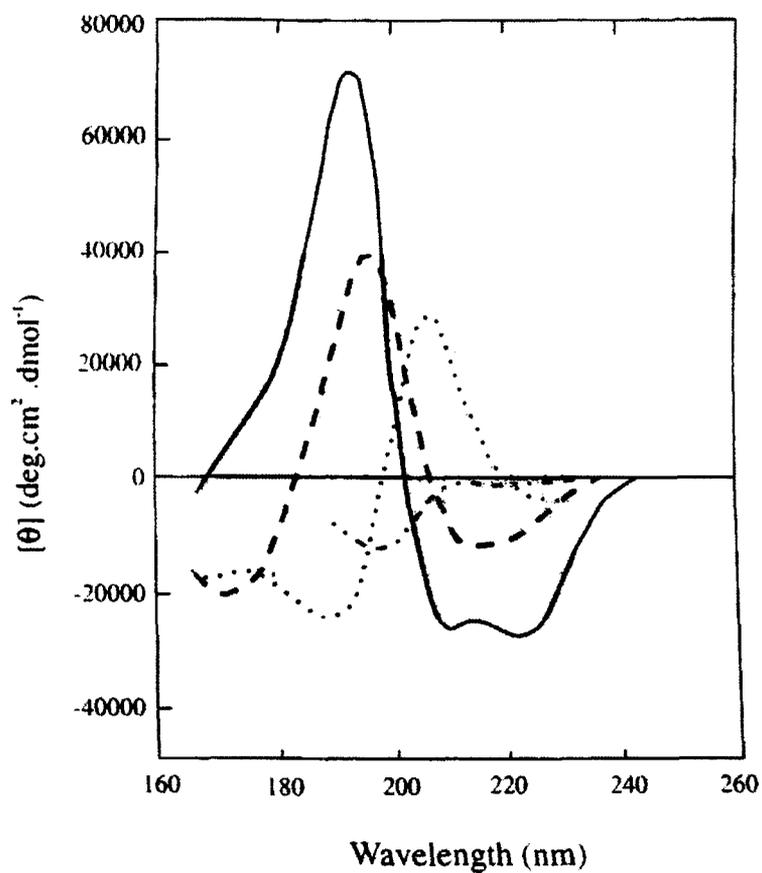
Currently, no reports of structural or functional research on Toc complex proteins have been done using fluorescence spectroscopy.

### **1.6.3 Circular dichroism spectroscopy**

Circular dichroism (CD) spectroscopy is a technique which measures the differential absorption of left and right circularly polarised light as it passes through chiral molecules (Kelly and Price, 2000). Briefly, plane-polarised radiation is passed through a modulator generally consisting of a piezoelectric quartz crystal tightly bound to an isotropic material (e.g. quartz). Alternating electric fields are applied to the crystal which causes structural changes to the crystal causing the quartz to transmit circularly polarised light at the extremes of the electric field (Kelly and Price, 2000). This polarised light is then passed through chiral molecules across a range of wavelengths and the resulting absorption of the left and right polarised light is recorded. As samples absorb the left and right circularly polarised light differently the resultant radiation is elliptically polarised (Kelly and Price, 2000). CD spectrometers, generally, detect the left and right light components separately and displays them either as the difference in absorbance ( $A$ ) of the two components or as the ellipticity in degrees ( $\theta$ ), where  $\theta = 32.98 * \Delta A$  (Kelly and Price, 2000). The resulting output is ellipticity plotted as a function of wavelength ( $\lambda$ ). In the case of proteins it is the peptide bond which absorbs the light in the far UV range (240nm to 190nm). Since the molar concentration of these bonds is important for subsequent analysis, the output is converted to molar ellipticity. Molar ellipticity, or mean residue ellipticity, is calculated as a function of the difference in molar absorbance (Kelly and Price, 2000). Figure 9 shows the typical far-UV spectra for  $\alpha$ -helix,  $\beta$ -sheet, type I  $\beta$ -turn, and random coil protein samples plotted as a function of molar ellipticity and wavelength. After data collection the next important step is the interpretation of the protein secondary structure.



**Figure 8.** Effect of tryptophan environment on emission spectra. Emission spectra 1-4 are apoazurin Pfl, ribonuclease T<sub>1</sub>, staphylococcal nuclease, and glucagon, respectively. Models of each protein are shown indicating the tryptophan locations in red and the schematic diagram to the right represents locations of tryptophan residues relative to the rest of the protein. Figure taken from Lakowicz, 2006.



**Figure 9.** Far UV spectra showing different secondary structures.  $\alpha$ -helix (solid line),  $\beta$ -sheet (dashed line), type I  $\beta$ -turn (dotted line), and random coil (dots and short dashes) are shown from wavelengths 170 nm to 240 nm. Figure taken from Kelly and Price, 2000.

Proteins are rarely composed of a single secondary structure type and thus adequate interpretation of proteins requires post data analysis. Interpretation of protein structure, which is made up of multiple types of secondary structure, utilize complex curve fitting approaches involving sets of spectra of structurally well characterised proteins (Sreerama and Woody, 2000). The CONTIN procedure, developed by Provencher and Glöckner (1981), was one of the first methods for estimating secondary structure composition. Today, web-based applications such as Dichroweb (<http://dichroweb.cryst.bbk.ac.uk>) use multiple methods containing well characterised proteins to automate secondary structure prediction.

To date, structural data has been collected for a number of the Toc components. AtToc75 has been shown to contain ~70%  $\beta$ -sheet with only 12%  $\alpha$ -helix when reconstituted into liposomes (Hinnah et al., 1997). Subsequent analysis of Toc75 was performed by calculating the hydrophathy index using a pattern database of known 3-D structures with mainly  $\beta$ -sheet revealing 16 membrane spanning domains (Hinnah et al., 1997). Recombinant Toc34 on the other hand has been shown to contain 34%  $\alpha$ -helical structure and 26%  $\beta$ -sheets/ $\beta$ -turns based on far-UV CD spectroscopic data (Schleiff et al., 2002). Crystal structures of pea Toc34 show similar results acquired from CD data revealing 36%  $\alpha$ -helical content and 18%  $\beta$ -sheet (Koenig et al., 2008). Most recently, the acidic domain of atToc159 and atToc132 have been shown to be 63% and 73% random coil, respectively (Richardson, 2008). It has been proposed that such an unstructured portion of the Toc159 proteins in *Arabidopsis thaliana* may contribute to the proteins' ability to recognize a large number of transit peptides with highly variable sequences. Further investigations into structural characteristics of Toc proteins are necessary in order to continue to develop our understanding of the translocation process into chloroplasts.

## **1.7 Objectives and hypotheses**

The overall objective for the current research is to investigate the interaction between Toc159 and Toc34 gene families in *Arabidopsis thaliana* using biophysical techniques. Research has shown that the interaction between these families of proteins is complex and difficult to study, however, has not been attempted using techniques such as fluorescence spectroscopy (FL) and circular dichroism spectroscopy (CD). An ultimate goal is to determine the binding affinities between members of the Toc159 and Toc34 protein families for the purpose of understanding how structurally distinct Toc complexes are formed. As an initial step toward this goal, aims of the current project are to elucidate physical properties of atToc159G and atToc33G that might be important for future studies of their interactions. Specific objectives are described below.

### **1.7.1 Observe interaction events between atToc159G and atToc33G using biophysical techniques**

The first objective is to establish that an interaction between atToc159G and atToc33G could be observed using biophysical (CD, FL,) and molecular biology techniques such as blue native polyacrylamide gel electrophoresis (BN-PAGE). Although the interaction between atToc159G and atToc33G has been observed using techniques such as pull-down assays and in a yeast two-hybrid system (Rahim et al., 2009) the interaction between these proteins had not yet been studied using biophysical techniques. Such biophysical techniques can identify characteristics of atToc159G/33G interactions, which can be quantified and potentially reveal structural changes as observed by changes to emission spectra given different experimental conditions. BN-PAGE provides a means to estimate the proportion of heterodimer formation. It has been observed that the formation of homodimers and heterodimers occurs, however, the proportions of homodimer:heterodimer have never been recorded (Sun et al., 2000; Koenig et al., 2008; Ivanova et al., 2004). Also, the creation of 3D models for atToc159G using alignment modelling software can reveal information about the tryptophan microenvironments critical for fluorescence emission spectra interpretation. It is hypothesised that the interaction between atToc159G and atToc33G

can be shown using FL and CD and confirmed using BN-PAGE. 3D models of atToc159G will also be generated from alignment models with atToc33G enabling us to predict the tertiary structure of atToc159G due to the moderate sequence identity between the two proteins (~36%; Weibel et al., 2003).

Objectives investigating the interaction between atToc159G and atToc33G also included the development of single tryptophan mutants. Creating mutant proteins can often give information about the effects that such mutations have on protein structure and function. Since intrinsic fluorescence is also used in the current research, eliminating a tryptophan residue from the proteins may also increase the sensitivity of the fluorescence assay. It is hoped that the single tryptophan mutants will generate important structural information about each tryptophan residue for both structure and function of atToc159G.

### **1.7.2 Manipulation of the monomer/dimer relationship between atToc159G and atToc33G**

The second objective was to manipulate the monomer/dimer ratio of each individual protein in an attempt to promote heterodimer formation and to establish the structural requirements of nucleotide and  $Mg^{2+}$  cofactor. This objective utilizes CD, FL, and BN-PAGE for analysis while introducing chelating agents in order to remove nucleotide from the proteins and promote monomer formation. Chelating agents remove  $Mg^{2+}$  cofactors from proteins, presumably causing the subsequent removal of nucleotide in the case of GTPases. Since GTP has been shown to be required for heterodimer formation, the removal of nucleotide and subsequent reintroduction of nucleotide in the presence of cofactor is predicted to prove useful for promoting the formation of heterodimers between atToc159G and atToc33G (Hiltbrunner et al., 2001). It is hypothesised that treatment with chelating agents will promote the formation of monomers for both atToc159G and atToc33G. Also, the introduction of GTP and  $Mg^{2+}$  ion in a controlled manner would promote the formation of heterodimers allowing for increased sensitivity for heterodimers detectable using both FL and CD.

## 2. Methods

### 2.1 Construction of wild-type and mutant atToc159G and atToc33G constructs

#### 2.1.1 Wild-type atToc33G and atToc159G constructs

Both atToc159G (At4g02510) and atToc33G (At1g02280) constructs were described previously (pET21d:atToc159G<sub>His</sub> & pET21d:atToc33G<sub>His</sub>) (Smith et al., 2002; Weibel et al., 2003). Coding sequences for each protein correspond as follows; atToc159G corresponds to basepairs 2180-3276 and atToc33G to basepairs 1-786 (Ivanova et al., 2004; Koenig et al., 2008). Each protein coding sequence was cloned into pET21-d, which generated an in-frame C-terminal hexa-histidine tag and introduced into *E. coli* strain BL21(DE3) to facilitate production of the corresponding recombinant proteins.

#### 2.1.2 Mutant atToc159G W973F and W1056F mutant constructs

AtToc159G single tryptophan mutants were constructed using an overlap-extension PCR-based method as previously described (Smith et al., 2002). Amino acids 727 – 1092 encode the G-domain, which contains two tryptophan residues located at amino acid residues 973 and 1056 (Figure 10; Ivanova et al., 2004). The two final constructs created were pET21d:atToc159G W973F and pET21d:atToc159G W1056F using the following method. For atToc159G W973F, the first round of PCR (95°C for 5 min, 94°C for 30 sec, 46°C for 30 sec, 72°C for 45 sec, repeated 39x; 72°C for 5 min, then held at 4°C until stopped) was done using pET21d:atToc159G<sub>His</sub> as the template as well as primer set #1 (appendix 1) and incorporated a 5' *NcoI* restriction site including a start codon and a codon for one additional alanine residue in order to maintain the reading frame, and the W973F mutation at the 3' end. The second round of PCR (all rounds of PCR used the same conditions as given above), using primer set #2 (appendix 1) and pET21d:atToc159G<sub>His</sub> as template, incorporated a 3' *XhoI* restriction site and the W973F mutation at the 5' end. The final PCR used the PCR products from both the first and second rounds of PCR as template

```

725  _TITSQDGTKLEFSMDRPAGLSSSLRPLKPAAAPRANRSNIFSNVNTMADE
772  TEINLSEEEKQKLEKIQSLRVKFLRLLQRLGHSÆDSIAAQVLYRLALLAG
820  RQAGQLFSLDAAKKKAVESEÆGNEELIFSLNILVLGKAGVGKSATINSIL
867  GNQIASIDAFGLSTTSVREISGFVNGVKITFIDTPGLKSAAMDQSTNAKML
917  SSVKVMKKCPDIVLYVDRLDIQTRDLNMLPLLRTITASLGPSIWKNAIV
963  TLTHAASAPPDGPGTPLSYDVFVAQCESHVQQSIGQAVGDLRLMNPILMN
1011 PVSLVENHPLCRKNREGVKVLENGQTWRSQLLLLCYSKVLSETNSLLRPQ
1057 EPLDHRKVFGERVR_

```

**Figure 10.** Amino acid sequence of atToc159G. The amino acid sequence for atToc159 GTP-binding domain (AA 727-1092). Tryptophan (W) locations are shown in red.

and primer set #3, creating the full length coding sequence for atToc159G including the W973F mutation as well as the restriction sites identified above. For atToc159G W1056F, a similar method was used to generate full-length coding sequence where the first and second rounds of PCR used primer set #4 and primer set #5, respectively (appendix 1). Full length PCR products for both the atToc159G W973F and atToc159G W1056F mutants were then cloned into pCR4BluntTOPO according to the manufacturer's instructions (Invitrogen). Each construct was then digested with the appropriate enzymes and ligated into the *NcoI* and *XhoI* restriction sites of pET21d generating pET21d:atToc159G W973F and pET21d:atToc159G W1056F with hexa-histidine tags on their C-termini. The new constructs were transformed into *E. coli* DH5 $\alpha$  and *E. coli* BL21 (DE3), and transformants were screened by restriction digests of isolated plasmids.

## **2.2 Expression and purification of recombinant wild-type atToc33G<sub>His</sub> and wild-type and mutant atToc159G<sub>His</sub> proteins**

### **2.2.1 Wild-type atToc159G<sub>His</sub> and atToc33G<sub>His</sub>**

Proteins were expressed in *E. coli* BL21(DE3) strains using the pET expression system (Novagen). Simply, 1 L LB broth containing 50  $\mu$ g/mL ampicillin was inoculated with 4 mL of overnight culture. Cultures were grown at 37°C, shaking at 240 RPM until OD<sub>600</sub> of 0.6 – 0.8 was reached. Protein expression was induced with 1 mM IPTG and incubated for 3 h at 37°C, shaking at 240 RPM. Cultures were then chilled on ice for 10 min before centrifugation at 6,000 g for 15 min (JLA.10.500, Beckman Coulter). The supernatant was discarded and the bacterial pellet was stored at -20°C until being processed further. Cell pellets were thawed on ice and resuspended in 50 mL of binding buffer (BB) (20 mM Tris-HCl, pH 7.9, 300 mM NaCl, 5 mM MgCl<sub>2</sub>, 10 mM imidazole, 1 mM Tris(hydroxypropyl)phosphine (THP)). Cells were lysed using a French press (Travenol Instrument Company) at a pressure of 11,000 psi with a flow rate of approximately 60 drops per minute. The

bacterial cell lysate was centrifuged at 75,600 g for 15 min to pellet insoluble material (JA.30.50, Beckman Coulter), and yield the total soluble extract in the supernatant.

Recombinant proteins were purified using immobilized metal affinity chromatography (IMAC). Briefly, a 2 mL column of Ni<sup>2+</sup>-charged NTA resin (Novagen) was washed with 10 column volumes of sterile milli-Q water, then with 10 column volumes of BB to equilibrate the column. The total soluble extract was applied to the column twice and the flow-through was collected. An initial wash of 10 column volumes of BB was applied to the column. A second wash was applied using 6 column volumes of wash buffer (WB) (20 mM Tris-HCl, pH 7.9, 300 mM NaCl, 5 mM MgCl<sub>2</sub>, 40 mM imidazole, 1 mM THP) to further remove any unbound or non-specifically bound protein. A final wash using 10 column volumes of sodium phosphate buffer (NaP) (50 mM NaPO<sub>4</sub>, 5 mM MgCl<sub>2</sub>, 1 mM THP) was applied to continue to remove any unbound or non-specifically bound proteins as well as perform a buffer exchange. Proteins were then eluted using elution buffer (EB) (50 mM NaPO<sub>4</sub>, 5 mM MgCl<sub>2</sub>, 1 mM THP, 250 mM imidazole) in six fractions of 1 mL into glass vials. Glycerol was added to a final concentration of 10%, and eluted proteins were stored at -80°C until analysed. Elution fractions were analyzed by resolving proteins using SDS-PAGE (4% stacking and 12% resolving) under constant current (15 mA during the stacking and 25 mA during the resolving) (Boyer, 2006). Gels were stained with Coomassie brilliant blue R-250 and destained using 10% acetic acid, 45% methanol, 45% water.

After analysis by SDS-PAGE, elutions containing significant concentrations of protein were dialyzed into CD buffer (50 mM NaPO<sub>4</sub>, 5 mM MgCl<sub>2</sub>, 1 mM DTT). Dialysis was done to remove fluorescent contaminants such as imidazole and THP. Three mL of protein was placed into a D-tube dialyzer MAXI (Novagen) and incubated in 400 mL of CD buffer for 3 h at 4°C. Protein was removed from the dialyzer and placed in clean glass vials and were either kept at 4°C for immediate analysis or returned to -80°C for longer-term storage, depending on if they were used for measurements immediately or on a subsequent

day. Proteins stored at -80°C for subsequent analysis were mixed with glycerol to a final concentration of 10% prior to freezing.

### **2.2.2 AtToc159G single tryptophan mutant proteins**

Similar to atToc159G, single tryptophan mutant proteins were expressed using the pET expression system (Novagen) with the following changes. Protein expression was induced with 2 g/L of The Inducer (Molecular, Columbia, Maryland, U.S.A.) and incubated for 16 h at 22°C, shaking at 125 RPM. Purification protocol for single tryptophan mutant atToc159G proteins was the same as for wild-type atToc159G. Also, purified proteins were dialysed into 400 mL of CD buffer for 3 h prior to circular dichroism and fluorescence spectroscopy measurements.

### **2.3 Nucleotide removal by chelating Mg<sup>2+</sup> cofactor**

Wild-type atToc159G and atToc33G proteins were treated with both EDTA and Chelex resin (Bio-rad) in order to chelate Mg<sup>2+</sup> ions bound to proteins in an effort to dislodge the nucleotide from the protein. For EDTA treatments, samples were treated with 100 mM EDTA for 1 hr at 4°C on a rotator before being applied to Zeba desalt spin columns (Thermo Scientific) as per manufacturer's protocol. The desalting step served to remove nucleotide, and Mg<sup>2+</sup> ions from the protein sample allowing for controlled introduction of such elements during measurements. Protein eluted from the Zeba column was used immediately for measurement.

Chelex resin is used for chelating metal ions while making recovery of sample easy. Samples were treated with 5 g / 100 mL for 3 h at 4°C on a rotator and then spun at 14,000 g for 10 min at 4°C. Samples were either used immediately for measurement or stored at 4°C with the Chelex resin for a maximum of 2 days before analysis.

## **2.4 Blue Native PAGE**

Blue Native PAGE was used to assess protein complexes in their natively folded state. Pre-cast NativePAGE™ Novex® Bis-Tris gels were purchased from Invitrogen with both 3 – 12% and 4 – 16% gradients. Gels were run at a constant voltage of 150 mV for 90 minutes. Proteins were fixed with destain (45% methanol, 45% water, 10% acetic acid) by microwaving on high for 30 seconds and incubating on a shaker for 30 minutes. Staining was done using Coomassie brilliant blue R-250 and destained until bands were visible.

## **2.5 Fluorescence Spectroscopy**

Fluorescence spectroscopy was performed using a Cary Eclipse fluorescence spectrophotometer (Varian, Palo Alto, CA, U.S.A.) with quartz cuvettes with a pathlength of 0.5 cm. Constant slit width (5 nm) for both excitation and emission was used with a scan rate of a 120 nm/min with an interval of 1 nm and a reading time of 0.5 sec. Excitation wavelength was 295 nm and chamber temperature was held constant at 22 °C. All spectra had buffer spectra subtracted before comparison and all concentrations and sample conditions are indicated in figures.

## **2.6 Circular Dichroism Spectroscopy**

Far-UV CD spectra were measured using an Aviv 215 spectropolarimeter (Aviv Biomedical). Measurements were performed using rectangular quartz cells with 0.1 cm pathlength. AtToc159G and atToc33G were measured in concentrations ranging from 2.5 μM to 5 μM as calculated by Bradford assay (Bradford, 1976). All measurements were done in CD buffer (Section 2.2.1) with added components where indicated, unless stated otherwise within figures. All samples were incubated at room temperature for 10 min prior to measurements. Spectra of samples and buffer were measured with a 0.5 nm/s scanning speed at 0.5 nm intervals, and are an average of at least eight scans and up to 16. All averaged spectra were buffer-subtracted using Aviv CDSD software (Aviv biomedical). Subsequent analysis

included smoothing and converting to mean residue molar ellipticity using the same software. Spectra were deconvoluted using the Dichroweb website using CDSSTR algorithms (Whitmore and Wallace 2004; Compton and Johnson 1986). For expected interaction spectra between atToc159G and atToc33G, individual spectra were subsequently converted to molar ellipticity.

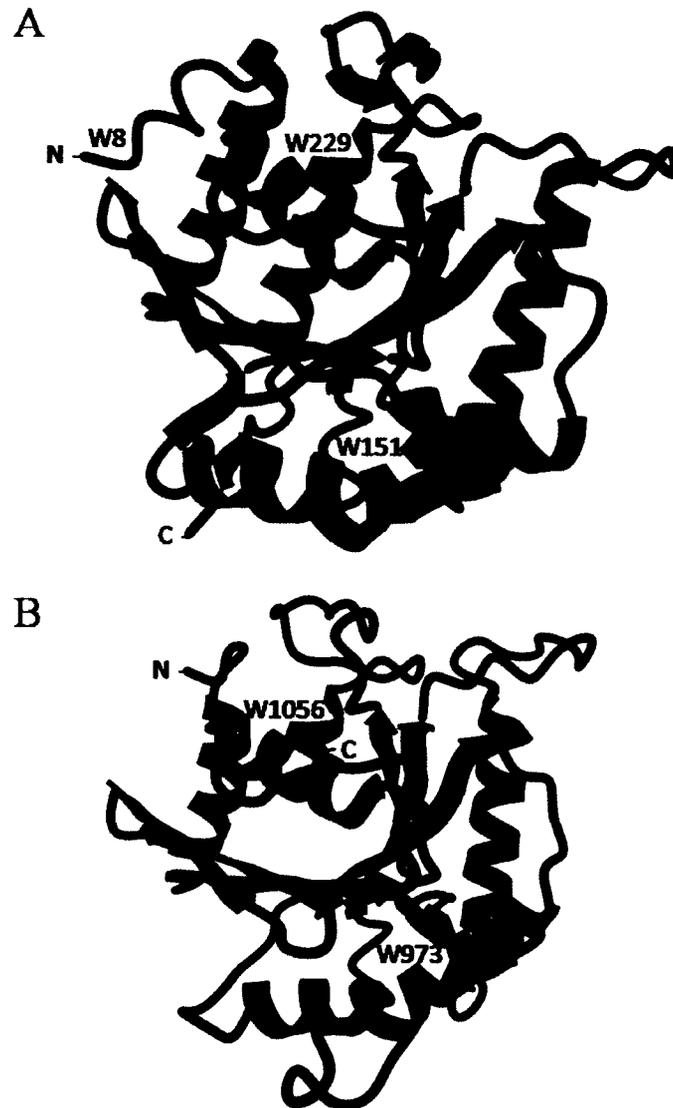
### **3. Results**

#### **3.1 Homology model of atToc159G**

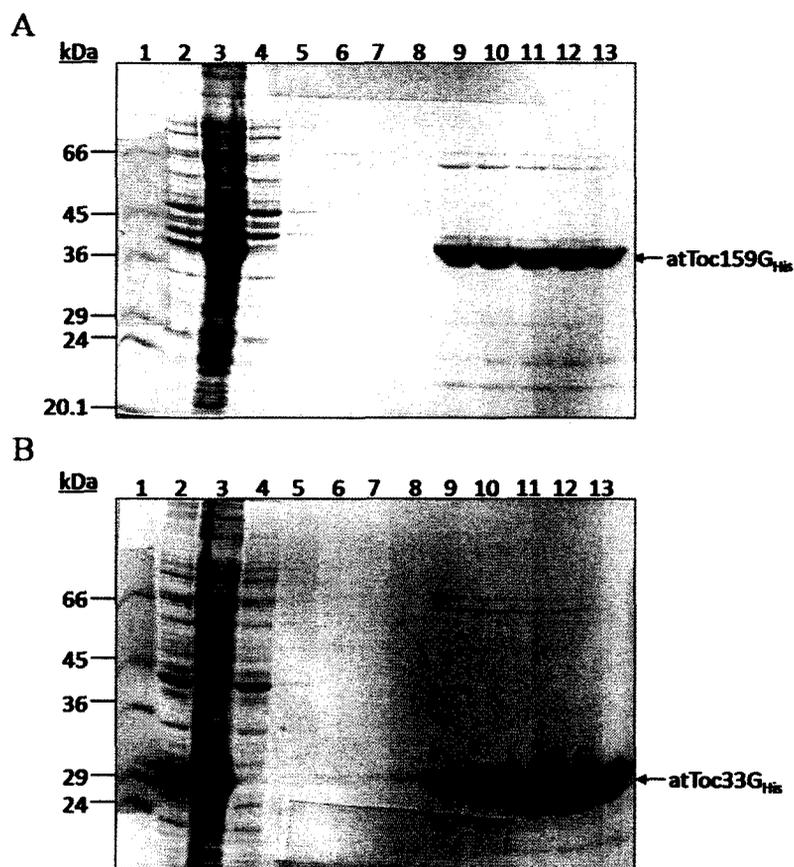
With the crystal structure of psToc34G and atToc33G recently elucidated in multiple nucleotide loading states, the utility of creating a homology model for atToc159G was appropriate (Sun et al., 2002; Yeh et al., 2007; Koenig et al., 2008). Using SWISS-MODEL from the Swiss institute for bioinformatics (<http://swissmodel.expasy.org/>), a homology model was generated based on the forced alignment with the atToc33G crystal structure (PDB 3BB3; Kiefer et al., 2009). The purpose of creating this homology model was to allow for comparisons of tryptophan microenvironments within 3-D space. Figure 11 shows both the 3D model of atToc33G (3BB3) and the homology model of atToc159G and their tryptophan locations (green). It is clear that two of the tryptophans in atToc33 are conserved in atToc159 and that these locations are relatively close to the surface of the protein which suggests an expected emission maxima of ~330nm (Figure 8). The homology model also suggests that atToc159G contains ~30%  $\alpha$ -helix and ~20%  $\beta$ -sheet, which is similar to the structural composition observed for atToc33G (Koenig et al., 2008). Also, it is noteworthy that tryptophan 229 in atToc33G and the comparable tryptophan in the atToc159G homology model (W1056) are close to the polyethylene glycol binding site observed by Koenig et al. (2008).

#### **3.2 Expression and purification of atToc159G, atToc33G, and atToc159G single tryptophan mutants**

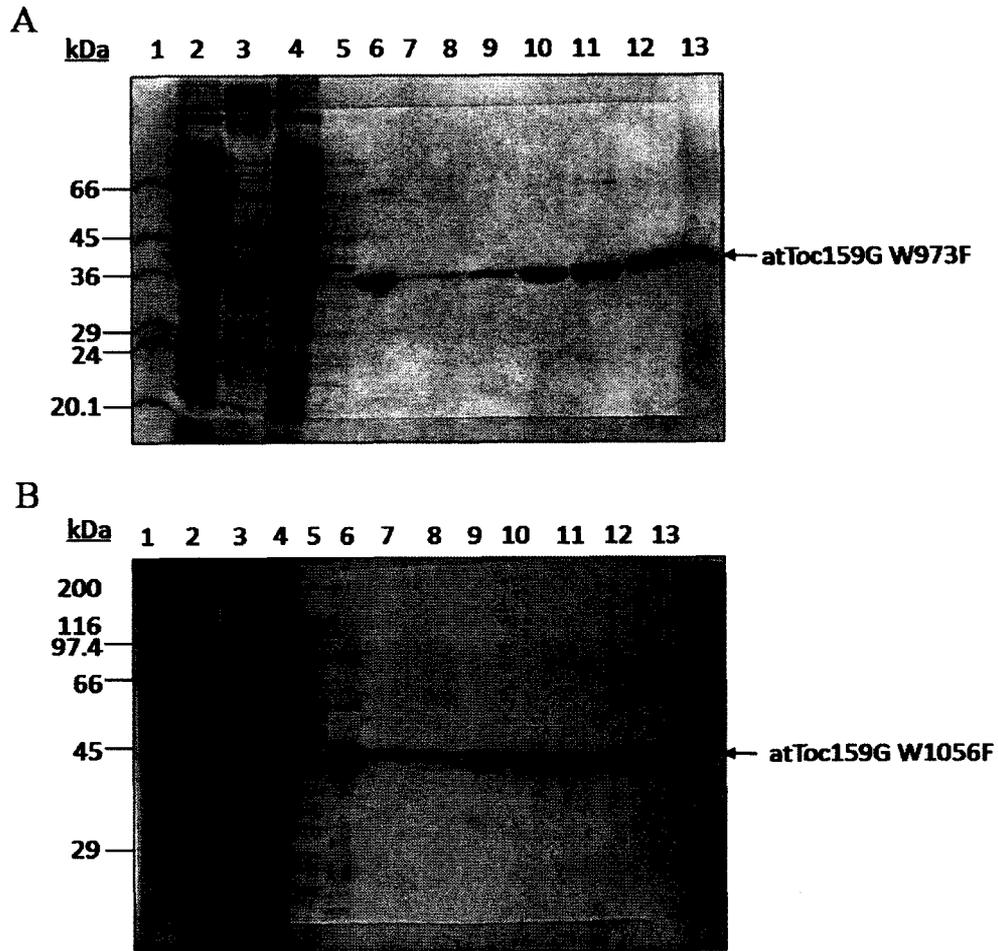
Histidine-tagged recombinant proteins were expressed in *E. coli* and purified using immobilized metal affinity chromatography (IMAC) for fluorescence and CD spectroscopy studies. Optimization of expression and purification conditions resulted in the effective purification of highly pure recombinant proteins as evidenced by the lack of higher or lower molecular weight bands in the elution fractions (Figure 12 and 13; lanes 8-13). AtToc159G wild-type and single tryptophan mutants are approximately



**Figure 11.** Crystal structure of the G-domain of atToc33 and homology model of atToc159G. Crystal structure of atToc33G (A; PDB 3BB3; Koenig et al., 2008) and homology model of atToc159G (B; Swiss model; Keifer et al., 2009) show significant amounts of similarity in overall structure as well as comparable tryptophan locations. Tryptophan locations are shown in green and labeled with residue position relative to the full length protein. N- and C-termini are also labeled. 3D models were generated using YASARA modeling software.



**Figure 12.** 12% SDS-PAGE of expression and purification of recombinant proteins atToc159G and atToc33G. (A) Purification of atToc159G (~42 kDa). Lane 1, molecular weight ladder; lane 2, soluble protein fraction, lane 3 pellet, lane 4 IMAC flowthrough, lane 5 BB wash, lane 6 WB wash, lane 7 NaP wash and lanes 8-13 are elution fractions 1-6. (B) Purification of atToc33G (~29 kDa) with same lane assignments as (A). Thirty  $\mu$ L of sample was loaded in each lane.

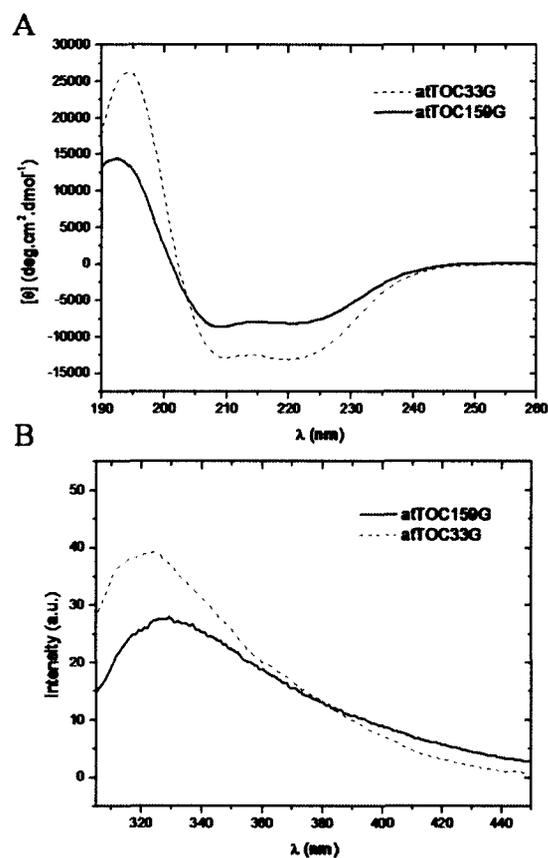


**Figure 13.** 12% SDS-PAGE of atToc159G single tryptophan mutants. (A) Purification of atToc159G W973F (~42 kDa). Lane 1, molecular weight ladder; lane 2 soluble protein fraction; lane 3 pellet, lane 4 IMAC flowthrough, lane 5 BB wash, lane 6 WB wash, lane 7 NaP wash and lanes 8-13 are elution fractions 1-6. (B) Purification of atToc159G W1056F (~42 kDa) with same lane assignments as (A). Thirty  $\mu$ L of sample was loaded in each lane.

42 kDa and atToc33G is approximately 29 kDa; their migration distances on SDS-PAGE gels are consistent with these molecular weights (Figs. 12 & 13). Unexpectedly, expression conditions for single tryptophan mutants required significant changes to the expression protocol when compared to wild-type proteins in order to optimize the amount of soluble protein harvested during purification. Preliminary results using IPTG induction resulted in expression; however, only as inclusion bodies within the *E. coli* (data not shown). An alternative inducing agent called The Inducer (Molecular, Columbia, Maryland, U.S.A.), which provides a lower affinity interaction than IPTG with the promoter, was used to induce expression of the mutants; this paired with lower induction temperatures, increased induction times, and lower shaking speeds allowed for the production of soluble proteins in relatively high yields (Figure 13; See Section 2.2.2).

### **3.3 Structure of wild-type atToc159G and atToc33G and their interaction**

Circular dichroism (CD) and fluorescence (FL) spectroscopy were used to observe the structural characteristics of recombinantly produced atToc159G and atToc33G. Results indicate that atToc159G contains about 25%  $\alpha$ -helix and 23%  $\beta$ -sheet whereas atToc33G contains 47%  $\alpha$ -helix and 11%  $\beta$ -sheet when deconvoluted using Dichroweb (Figure 14A; Whitmore and Wallace 2004). As compared to recent crystal structures generated for atToc33G, similar secondary structure characteristics can be observed from the CD deconvolution data; however, since no structural data has been reported for atToc159G only comparison to the homology model can be made (PDB 3BB3; Koenig et al., 2008). The homology model of atToc159G predicts ~30%  $\alpha$ -helix and ~20%  $\beta$ -sheet which are comparable to the data derived from deconvolution giving confidence to both the homology model as well as the folded state of the recombinant protein. FL emission spectra identify slightly different tryptophan environments for atToc159G and atToc33G as revealed by the  $\lambda_{\text{max}}$  for each protein (Figure 14B; 159G  $\lambda_{\text{max}} = 330\text{nm}$ ; 33G  $\lambda_{\text{max}} = 325\text{nm}$ ). However, these differences are likely to be a result, at least in part, of atToc33G having

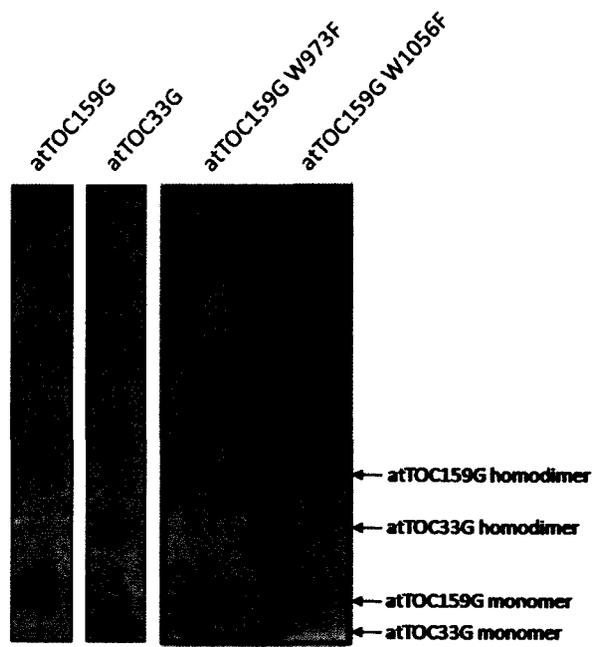


**Figure 14.** CD and Fluorescence spectra of atToc159G (5  $\mu\text{M}$ ) and atToc33G (5  $\mu\text{M}$ ) at 20°C and pH 7.5. (A) CD spectra shows that atToc159G and atToc33G both have secondary structures containing both  $\alpha$ -helix and  $\beta$ -sheet as is expected based on previous crystal structures of atToc33G (Koenig et al., 2008). (B) Tryptophan fluorescence shown for both wild-type proteins is typical for residues located on the interior of the secondary structure of the protein (Lakowicz, 2006). Note that atToc159G contains two tryptophan residues while atToc33G contains three.

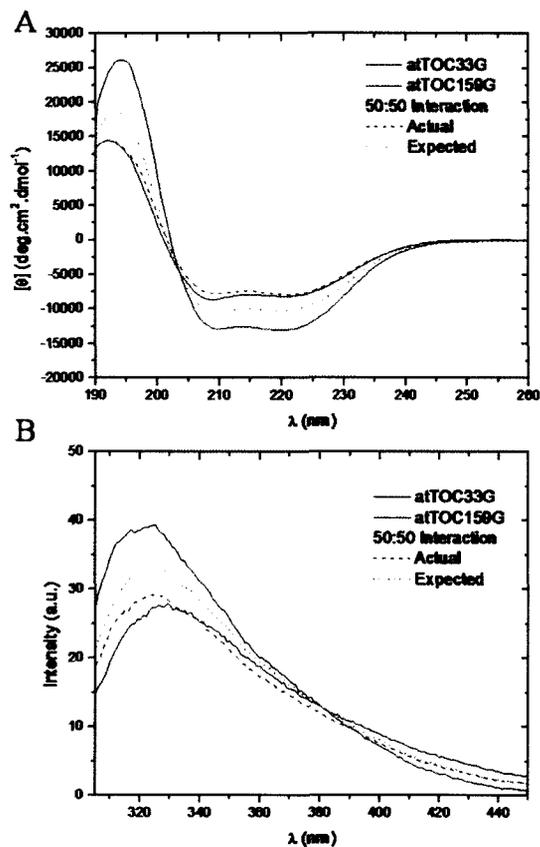
three tryptophan residues as compared to two for atToc159G; thus, the overall tryptophan environments for these proteins are not expected to be exactly the same. Nonetheless, the blue shifted emission intensity maxima of each protein may be a result of the location of the tryptophan residues within the tertiary structure of each protein as shown in 3-D models (Figure 11).

Blue native PAGE (BN-PAGE) shows proteins resolved on a gel in their natively folded (oligomeric, if applicable) state (Schagger and von Jagow, 1991). Previous observations of homodimer formation of atToc33G *in vitro* are in agreement with analysis of the protein using BN-PAGE, which shows the presence of both monomer and dimer (Figure 15; Yeh et al., 2007). AtToc33G and atToc159G resolved using BN-PAGE indicate the presence of both monomeric and homodimeric forms of the proteins with approximately a 3:1 monomer:dimer ratio for atToc33G and a 2:1 ratio for atToc159G as calculated using ImageJ software (appendix 2; NIH; Abramoff et al., 2004). These results support previous findings using gel filtration which observed recombinant atToc33G in a predominantly monomer state (Yeh et al., 2007).

Early objectives of the current project focused on investigating the interactions between the G-domains of Toc159 and Toc34 gene families. Mixing the proteins together and assessing spectral changes was expected to give information of interaction events and thus proteins were mixed in a 1:1 dilution. Interaction events can cause changes to the overall CD spectra as structural/conformational changes often occur during such events (Kelly and Price, 2000). In the event of no interaction, the addition of individual spectra would be expected to give a similar result as that observed experimentally. Similarly, FL gives information about the microenvironment of tryptophan and changes to such environments can result in changes in the emission spectra and may be indicative of an interaction event (Lakowicz, 2006). Again, no interaction between proteins would result in a spectra similar to the addition of the emission spectra of each individual protein. Figure 16 shows the experimental and expected spectra for the interaction between atToc159G and atToc33G for both CD (A) and FL (B) data. CD results (Figure 16A)



**Figure 15.** Blue native PAGE of wild-type atToc33G and wild type and mutant atToc159G. Each individual protein was run in its native state allowing identification of monomers and homodimers. The atToc159G W1056F protein shows multiple bands possibly due to the formation of higher order oligomers or aggregates.



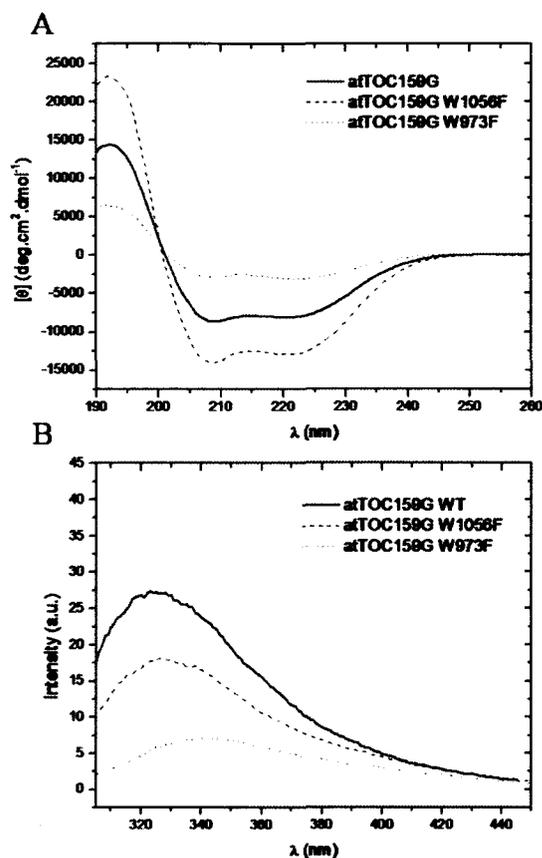
**Figure 16.** CD and fluorescence spectra showing atToc159G and atToc33G interaction. (A) Wild-type CD spectra as well as experimental and expected interaction spectra are shown. Overall shape of the interaction spectra remains similar to expected results however significant changes to intensity resulted. (B) Fluorescence spectra show similar results to CD data identifying a common  $\lambda_{\text{max}}$  between experimental and expected results with changes only in intensity. Results indicate an interaction between the proteins can be observed.

reveal a change in intensity when compared to the expected spectra potentially identifying an interaction event. Similar results are shown by FL (Figure 16B) showing a lower intensity than that expected if no interaction again revealing an interaction event between atToc159G and atToc33G. Confirmation of the interaction was visualized using BN-PAGE. Increasing amounts of either atToc159G or atToc33G added to their counterpart shows an increase in heterodimer formation (Figure 17). In the case of 5  $\mu$ M of each atToc159G and atToc33G, the BN-PAGE gel shows a band located between the homodimer bands for each protein. This intermediate-sized band is indicative of the atToc159G-atToc33G heterodimer. However, when compared to the total intensity of all bands in the sample, the heterodimer accounts for only 4.5% (appendix 2; value estimated using ImageJ) of the total load, which may partially explain the difficulty in detecting this interaction experimentally.

### **3.4 Structure of atToc159G single tryptophan mutants**

Mutants can often give information about the importance of a given residue for generating structure necessary for function. For example, the atToc33G R130A mutation shows a loss in dimer formation which has been reported to result in a decrease in GTP hydrolysis, however, more recent results dispute this loss in hydrolytic activity indicating the mutation's role in dimer formation only (Yeh et al., 2007; Koenig et al., 2008b; Lee et al., 2009). The W1056F and W973F mutations were characterized structurally in this study using CD and FL as well as with BN-PAGE. For atToc159G W973F the results show no changes to monomer/dimer proportions with respect to wild-type when observed using BN-PAGE (Figure 15). Significant changes in fluorescence emission intensity were, however, observed as was a shift in  $\lambda_{\text{max}}$  towards 350 nm due to the removal of tryptophan 973 (Figure 18B). Significant changes to CD spectra reveal a major loss in  $\alpha$ -helical structure from 25% to 4% and a loss of  $\beta$ -sheet from 23% to 15% when deconvoluted using dichroweb (Figure 18A). This loss in structure of atToc159G W973F corresponds to a significant loss in dimer formation. As indicated by BN-PAGE, the monmer:dimer ratio changed from 3:1 to 7:1 when calculated using ImageJ (Figure 15; appendix 2).



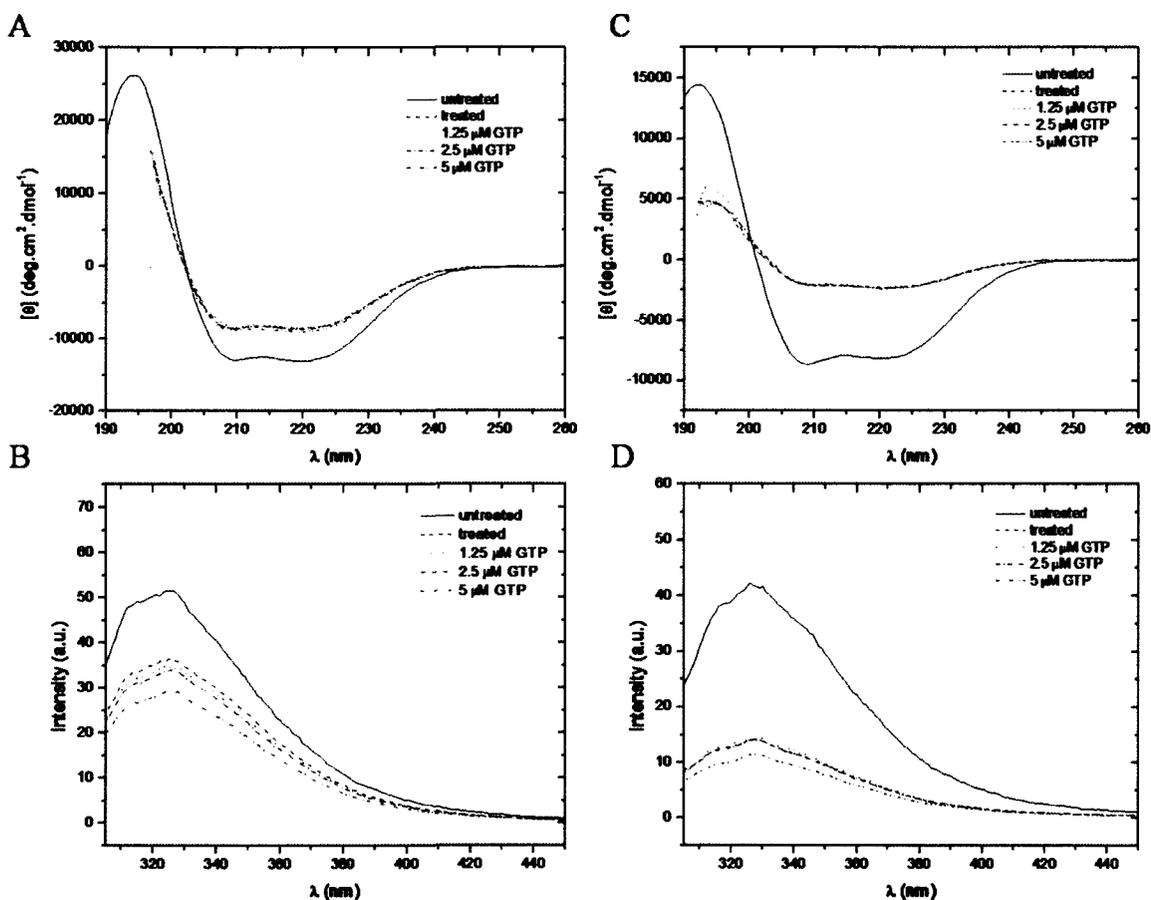


**Figure 18.** CD and Fluorescence spectra of atToc159G single tryptophan mutants (5  $\mu$ M) at 20°C and pH 7.5. (A) Comparison of the spectra of wild-type and mutant proteins reveals major structural changes as a result of mutations. (B) Tryptophan fluorescence shows a decrease in intensity for both mutants due to the loss of one tryptophan. The W973F mutation appears to cause a significant loss of structure resulting in a major loss in fluorescent intensity as well as a shift in  $\lambda_{max}$  towards higher wavelengths possibly due to the formation of aggregates. The W1056F mutant maintains a predominantly  $\alpha$ -helical structure with some  $\beta$ -sheet and similar  $\lambda_{max}$  when compared to wild-type atToc159G.

Collectively, the data suggest that the W973F mutation causes significant changes to the atToc159 G-domain, which may render the protein inactive. The atToc159G W1056F mutation also results in major structural changes observed by BN-PAGE, CD, and FL (Figure 15 and Figure 18). A shift in structure from 25% to 44%  $\alpha$ -helix and 23% to 11%  $\beta$ -sheet identifies a gain in secondary structure and the possibility of aggregation of the protein is revealed by BN-PAGE (i.e. appearance of higher-order oligomers) (Fig. 15). The fluorescence spectra show a loss of intensity resulting from the removal of a tryptophan residue. The formation of monomers, dimers, and multimers observed in the BN-PAGE supports the idea of aggregate formation indicating the importance of tryptophan 1056 for structure. Tryptophan 1056 is also shown to reside close to the putative transit peptide binding site observed by Koenig et al. (2008), possibly revealing its requirement for structure suitable for binding to transit peptides or other Toc proteins.

### **3.5 EDTA and Chelex treatment of wild-type proteins atToc159G and atToc33G**

Nucleotide binding proteins are known to have essential cofactors such as  $Mg^{2+}$  required for nucleotide binding and guanine nucleotide exchange (Zhang et al., 2000). In the case of the Toc-GTPases, cofactors are essential for the binding of GTP and GDP. Recent crystal structure evidence reveals that divalent Mg ions are present in all nucleotide loading states for Toc34 proteins in both *Pisum sativum* and *Arabidopsis thaliana* (Sun et al., 2002; Yeh et al., 2007; Koenig et al., 2008). If  $Mg^{2+}$  is a requirement for nucleotide binding, then treatment with chelating agents like ethylenediaminetetraacetic acid (EDTA) and Chelex should help dislodge nucleotides from GTP-binding proteins by chelating the ionic cofactor allowing for assessment of structural changes as a result of no bound nucleotide. CD, FL, and BN-PAGE were used to assess the structural changes that resulted from treatment of Toc159G and Toc33G with chelating agents EDTA and Chelex. Treatment of atToc159G and atToc33G with EDTA resulted in significant structural changes for both proteins that are apparently irreversible with the addition of nucleotide and  $Mg^{2+}$ , as observed by CD (Figure 19 A and C). Major losses in fluorescence

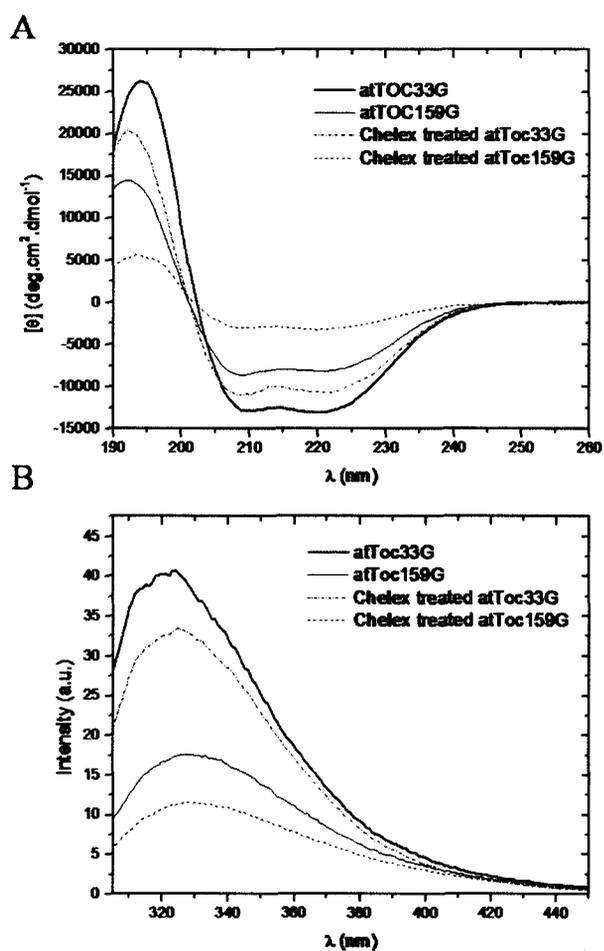


**Figure 19.** EDTA treatment of 5  $\mu\text{M}$  atToc159G and atToc33G. AtToc33G (A) and atToc159G (C) CD spectra after treatment with EDTA and the addition of 5  $\mu\text{M}$   $Mg^{2+}$  and increasing amounts of GTP added back in. (B) and (D) represent the corresponding fluorescence spectra ( $\text{Ex} = 295\text{nm}$ ) for both atToc33G and atToc159G, respectively.

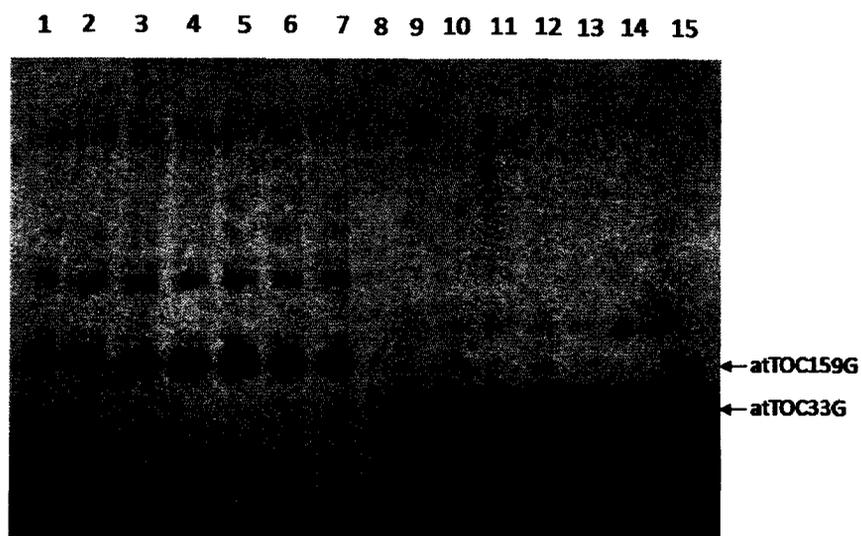
intensity were also observed for treated proteins despite the re-addition of  $Mg^{2+}$  and GTP back to the proteins, and are also indicative of structural changes (Figure 19 B and D). Due to its ability to interact with divalent cations, EDTA has the potential to maintain low-affinity bonds with proteins (Cui et al., 1994). It was expected that passing the EDTA-treated proteins over a desalting column would eliminate the free  $Mg^{2+}$ , nucleotide, and EDTA from the solution; however, bound EDTA is suspected to have been carried through the column with the protein. Structural changes during reintroduction of nucleotide and cofactor may then have been fruitless as the protein-bound EDTA would be available to chelate the added  $Mg^{2+}$  needed for nucleotide binding. An alternative to EDTA that does not bind to proteins and is therefore easier to remove was used to confirm the observations made on the EDTA-treated proteins. Such an alternative is Chelex resin.

Chelex resin is produced solely by Bio-Rad and allows for simple and efficient chelation of ions with a high capacity resin. The resin also simplifies the recovery of sample and lowers the overall time spent preparing proteins. Again CD, FL, and BN-PAGE were used to assess the structural changes that resulted from treatment of atToc159G and atToc33G with Chelex resin. The results indicate that the treatment with Chelex resin achieved similar results when compared to the EDTA-treated proteins for both CD and FL with respect to the effect of chelating agents on atToc159G and atToc33G (Figure 19 and Figure 20). Such structural losses of each protein as observed by CD show the necessity of bound nucleotide for maintaining structure (Figure 20A). Similarly, the loss of fluorescence intensity for both atToc159G and atToc33G was observed; however, the degree of loss of intensity was less when compared to EDTA treatment (Figure 20B). BN-PAGE of chelex-treated proteins also reveals a loss of dimer formation, shifting the monomer:dimer ratio from 3:1 to approximately 10:1 when compared to untreated wild-type proteins (appendix 2; Figure 21).

After characterizing the structural changes that occur during chelex treatment, nucleotide and  $Mg^{2+}$  cofactor were reintroduced in an attempt to reform a similar structure to untreated proteins. Results

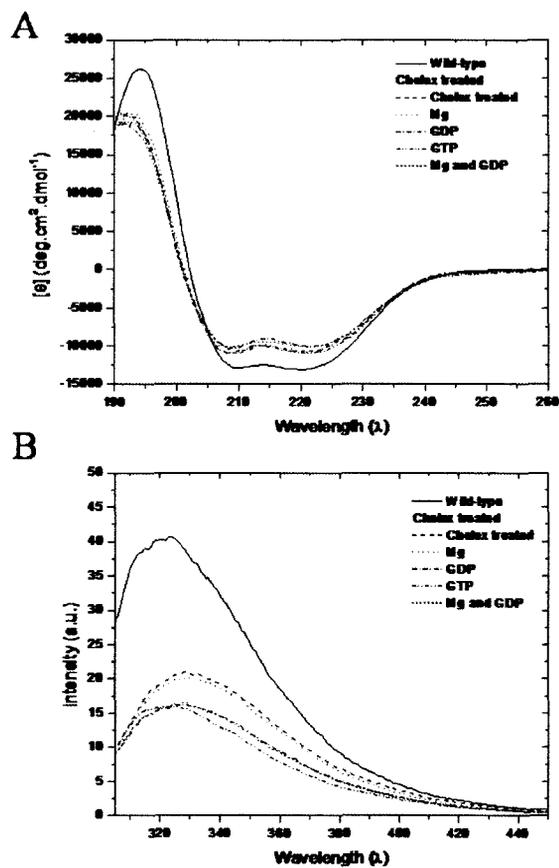


**Figure 20.** CD and FL spectra of 2.5  $\mu\text{M}$  wild-type and chelex treated atToc159G and atToc33G. CD spectra (A) and FL emission spectra (Ex = 295 nm; B) for atToc159G and atToc33G treated with chelating agent Chelex resin. Untreated protein spectra are shown with solid lines and treated protein spectra are shown with dashed (atToc33G) and dotted (atToc159G) lines.

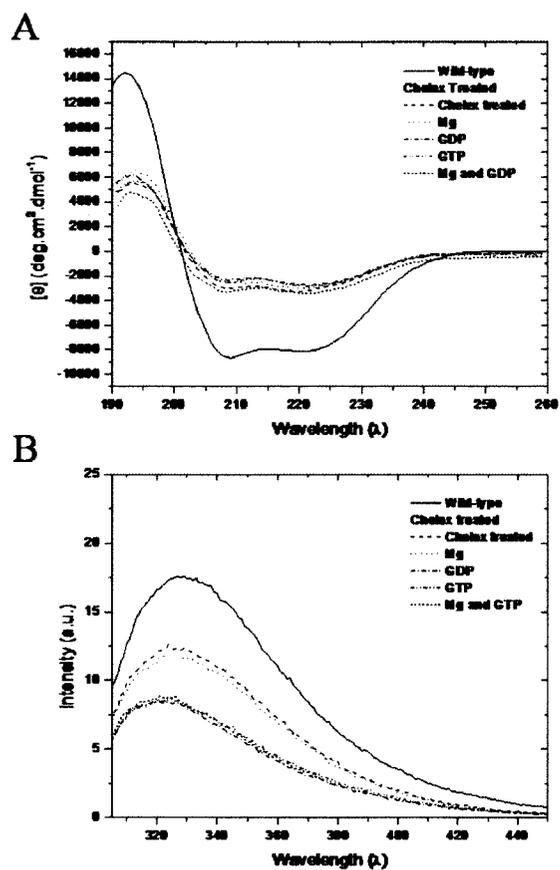


**Figure 21.** BN-PAGE of Chelex-treated atToc159G and atToc33G. Chelex treated atToc159G (Lanes 1-7) and atToc33G (Lanes 9-14) proteins with increasing amounts of  $Mg^{2+}$  added. Lane 1/9, untreated protein; lane 2/10, chelex treated protein; lane 3/11, chelex treated protein with  $1 \mu M Mg^{2+}$ ; lane 4/12, chelex treated protein with  $2 \mu M Mg^{2+}$ ; lane 5/13, chelex treated protein with  $4 \mu M Mg^{2+}$ ; lane 6/14, chelex treated protein with  $8 \mu M Mg^{2+}$ ; lane 7/15, chelex treated protein with  $10 \mu M Mg^{2+}$ . Lane 8 is empty.

show that chelex-treated atToc33G is unable to reform structure in the presence of nucleotide and  $Mg^{2+}$  indicating the necessity for a constant nucleotide loaded state, at least *in vitro*. CD and FL data for atToc33G shows that chelex treatment results in structural changes, and that the addition of  $Mg^{2+}$ , GDP, or GTP does not elicit any structural changes or changes to fluorescence emission spectra (Figure 22). More specifically, there is no change towards pre-treated protein behaviour with respect to fluorescence intensity and CD signal intensity. The further decrease in fluorescence emission intensity observed during the addition of nucleotide is likely a result of the absorption of the nucleotide itself (data not shown). AtToc159G FL and CD results reveal similar data as with atToc33G where no change was observed upon the addition of  $Mg^{2+}$ , GDP, or GTP to Chelex-treated protein (Figure 23). In one instance, the addition of  $Mg^{2+}$  and GTP did elicit a structural change towards wild-type (data not shown), however, such data has not been confirmed at this time.



**Figure 22.** Effects of Mg<sup>2+</sup> and nucleotide on Chelex-treated atToc33G. (A) Far UV CD spectra of 2.5 μM atToc33G pre- and post- chelex treatment as well as the addition of Mg<sup>2+</sup>, GDP, and GTP to chelex treated proteins. (B) FL emission spectra (Ex = 295 nm) for atToc33G with same treatments as A.



**Figure 23.** Effects of  $\text{Mg}^{2+}$  and nucleotide on Chelex-treated atToc159G. (A) Far UV CD spectra of 2.5  $\mu\text{M}$  atToc159G pre- and post- chelex treatment as well as the addition of  $\text{Mg}^{2+}$ , GDP, and GTP to chelex treated proteins. (B) FL emission spectra (Ex = 295 nm) for atToc159G with same treatments as A.

## 4. Discussion

The primary objectives of the current research were to investigate the interaction between atToc159G and atToc33G using biophysical and molecular techniques. Specific aims were to use FL, CD, and BN-PAGE to characterize the occurrence of both homodimeric and heterodimeric structures. It was also intended to characterize a series of single tryptophan mutants of atToc159G to assess their utility in such studies. Chelating agents were also used in an attempt to promote the formation of monomers, with the hope of subsequently promoting the formation of heterodimers. Results indicate the formation of both homodimeric and heterodimeric structures, as well as major structural changes as a result of tryptophan mutations. How these data contribute to the current research regarding chloroplast protein translocation is discussed below.

### 4.1 Homodimerization of atToc33G and atToc159G

One aim of this study was to investigate the occurrence of homodimers of both atToc159G and atToc33G. The overall secondary structure of purified recombinant proteins, as observed in this study using far-UV CD, agrees with previous findings for atToc33G (Koenig et al., 2008) and the homology model for atToc159G (Figure 11), establishing credibility of the recombinantly produced proteins; however, functional assays were not done to more definitively confirm proper protein folding. In the current study, it has been demonstrated that both atToc159G and atToc33G exist in a predominantly monomeric state as shown by BN-PAGE (Figure 15, Appendix 2), which is consistent with previous reports (Yeh et al., 2007; Reddick et al., 2007). It has also been shown that the monomer/dimer ratio for pea Toc34 is concentration dependent and is shifted towards the dimer at concentrations higher than 12.5  $\mu\text{M}$  (Reddick et al., 2007). Thus, it should be noted that proteins were analyzed at concentrations of approximately 5  $\mu\text{M}$ , and that the lack of concentration-dependent studies are a limitation of the current research. The dominant monomeric conformation of both atToc159G (~66%) and atToc33G (~75%) observed in the current study may be attributed to the equilibrium established by each individual protein

at concentrations of 5  $\mu\text{M}$  (Appendix 2). This equilibrium may reflect the protein pool in which adequate 'free' or monomeric protein is available for association with other Toc components in and around the time of translocation. It is not known whether transit peptides associate with Toc GTPases in either the monomeric or dimeric form. It has been proposed that homodimeric structures play a role in recognition of transit peptides but this has not been confirmed (Li et al., 2007). It is plausible that this high proportion of monomeric protein has a greater potential to interact with transit peptides in preparation for a translocation event. This increased pool of monomeric protein may increase the efficiency of transit peptide-Toc protein interactions, which may be necessary for the efficient formation of Tic/Toc supercomplexes (Chen and Li, 2007). The concentration-dependent dimerization observed by Reddick et al. (2007) may also be relevant as localized protein concentrations close to the sites of complex formation may promote dimerization while monomeric forms maintain wide spread coverage of the chloroplast OEM in order to interact with transit peptides.

Secondly, and possibly more intriguing, homodimers exist for both atToc159G and atToc33G. Recent crystal structures, native gel electrophoresis and gel filtration studies have shown the homodimeric structure of atToc33G and propose its importance during protein translocation into chloroplasts and effects on GTP hydrolysis rates (Sun et al., 2002; Weibel et al., 2003; Li et al., 2007; Yeh et al., 2007; Koenig et al., 2008; Agne et al., 2009; Lee et al., 2009). Original investigations proposed that the dimer partner acted as a coGAP for GTP hydrolysis; however, more recent data shows that changes to hydrolysis rates are not suggestive of any coGAP functionality and maintain that homodimerization may either require a third protein to act as the coGAP or that it is the monomeric state that is required for functional translocation (Koenig et al., 2008b). Current data shows a significant proportion of both atToc159G and atToc33G are present as homodimers in solution, and remain as such when they are mixed, rather than reorganizing into heterodimers (Figure 17). However, it should be noted that no concentration-dependent experiments were attempted and is a known limitation to the current research. Treatments with Chelex significantly shifted the monomer:dimer ratio further toward the monomeric

state, suggesting the need for nucleotide and  $Mg^{2+}$  for dimer formation (Figure 21; Appendix 2). Also, stoichiometric studies have all agreed that there is only one Toc159 protein present in Toc complexes (Kikuchi et al., 2006, Schleiff et al., 2003; & Vojta et al., 2004). Current data shows a significant proportion of recombinant atToc159G exists as a dimer, suggesting that there may be some role atToc159 homodimerization plays in and around translocation. However, it is not clear whether atToc159 homodimerization occurs prior to Toc complex formation and may only contribute during transit peptide binding or nucleotide exchange. Another possibility is that the production of recombinant atToc159G produces homodimers as a result of the purification process; however, this factor should then be a consideration in interpreting all *in vitro* interaction data for Toc34 as well. Nonetheless, the *in vivo* conformation of both Toc159 and Toc34 protein monomers and their roles in translocation as well as evidence for homodimeric importance still remains to be fully explained for both atToc33G and atToc159G.

#### **4.2 Heterodimer formation is an uncommon event *in vitro***

The interaction between the Toc159 and Toc34 gene families has been previously observed and has been shown to be mediated by the GTP-binding domains of each protein (Smith et al., 2002). Currently, the formation of heterodimers has been observed using BN-PAGE as well as FL and CD for the first time (Figures 16 & 17). The formation of stable heterodimeric structures between atToc159G and atToc33G has important implications for current hypotheses. All current models for translocation include the interaction between Toc159 and Toc34 gene families including preferential interactions between members of each gene family producing functionally distinct Toc complexes. It has also been previously thought that the formation of heterodimers is a transient state *in vivo* and cannot be quantified (Li et al., 2007); however, recent studies and current observations show a stable formation of heterodimers produced from atToc159G and atToc33G both *in vivo* and *in vitro* (Rahim et al., 2008, Figure 17). The importance of homodimeric interactions has also been shown *in vivo* using atToc33

knockout mutants where complementation with atToc33 mutants unable to form homodimers inhibits the rate of translocation but not the affinity of preproteins for Toc receptors (Lee et al., 2009). This data contributes to the notion that the formation of heterodimers between atToc159 and atToc33 is important for translocation. Currently, it has been shown by BN-PAGE that the interaction between atToc159G and atToc33G at 5  $\mu$ M is only observed as approximately 4.5% of the total protein load consistent with previous findings (Appendix 2; Smith et al., 2002). Also, the interaction appears to be detectable using FL and CD; however, differences are minimal, possibly due to the low levels of heterodimer formation. Previous research revealed an overall stoichiometry of approximately 1:3:3 for Toc159:Toc75:Toc33 of the Toc complex (Kikuchi et al., 2006). In the current study, the G-domains of both atToc159G and atToc33G did not show any signs of higher order oligomers when analysed individually (Figure 15), however, a single band showing a complex larger than that of the homodimer for atToc159G when proteins were mixed may reveal an interaction of a combination of multiple atToc159G's and atToc33G's (Figure 17). This may be an indication of multiple interactions between atToc159G and atToc33G responsible for the observed complex sizes in previous studies (Kikuchi et al., 2006). To remain consistent with proposed stoichiometries, it is hypothesized that a single atToc159G may interact with up to four atToc33G's to form the complexes observed using BN-PAGE. It is also assumed by this data that the G-domains of atToc159 and atToc33 play an important role in Toc complex assembly. Confirmation of previous observations for heterodimer formation is needed as interactions may also involve other domains, such as the A-domain of atToc159 or the TM-domains of the Toc34 proteins. Studies utilizing two-dimensional gel electrophoresis may be effective in identifying other interactions between Toc proteins as initial native PAGE would maintain the interactions between proteins and secondary SDS-PAGE would identify individual components of the interaction.

### 4.3 AtToc159G tryptophan mutations

As a secondary objective, two single tryptophan mutants of atToc159G were generated in an attempt to increase the sensitivity of the fluorescence spectroscopic assays by limiting the number of tryptophan residues present. The recombinant proteins produced by the single tryptophan mutants revealed major structural changes as observed by CD spectroscopy making them inappropriate for interaction studies with atToc33G. Protein structure is essential for function and it has been demonstrated previously and in the current study that even changing a single residue can alter the structure of a protein significantly resulting in a loss of function (Petsko and Ringe, 2006). Modeling of atToc159G using SWISS-MODEL allowed for the identification of predicted tryptophan locations within the 3D structure of the protein. AtToc159G tryptophan mutant W973F results in a protein with significant losses in secondary structure and a loss of fluorescence intensity. Modeling predicted that this residue exists away from the putative transit peptide binding site and near the surface of the protein. Potentially, changes to the residue may indirectly modify the nucleotide binding site making it unable to bind nucleotide. Similar structural changes are seen for both EDTA-treated and Chelex-treated atToc159G suggesting that the W973F mutation may result in an inability of the protein to bind Mg and nucleotide. The W1056F mutation reveals an opposite effect with respect to secondary structure when compared to the W973F mutation. A marked increase in secondary structure is observed with only a moderate loss of fluorescence intensity consistent with removal of a single tryptophan (Fig. 18). The tryptophan W1056F mutation results in an increase in  $\alpha$ -helical content (25%  $\rightarrow$  44%) and a decrease in  $\beta$ -sheet (23%  $\rightarrow$  11%) which may be indicative of changes to the central  $\beta$ -basket which results in the  $\alpha$ -helical content increase. Also, results indicate that higher order oligomers occur upon mutation of W1056, which appears to be a result of changes to the surface of the protein in which an alternative protein-protein interaction site arises. Structural changes to the protein may not be significant as the fluorescence data indicates a similar tryptophan environment for both WT and W1056F atToc159G's. It is hypothesized that W1056F, which is located in the putative transit peptide binding site proposed by Koenig et al. (2008), may act to help

regulate the protein-protein interactions within the binding site and the current mutation has resulted in a loss of such regulation.

#### **4.4 Future investigations**

It has been shown that the interaction between atToc159G and atToc33G can be visualized using native gel electrophoresis; however, the proportion of proteins which exist in the heterodimeric interaction is low and difficult to quantify. Attempts to shift the proportions of monomer and dimer in an effort to subsequently promote heterodimer formation proved difficult as nucleotide removal irreversibly modified the structure of proteins. Novel techniques for promoting the detection of heterodimer formation would prove valuable to extending current data. Previous research has identified the formation of structurally distinct and functionally distinct Toc complexes (Ivanova, et al., 2004; Kubis et al., 2004; Smith et al., 2004). Developing a technique that would allow for the quantification of heterodimeric interactions would prove invaluable for the testing of current hypotheses. Techniques such as co-expression in *E. coli*, yeast two-hybrid and fluorescence resonance energy transfer may provide insight into the interaction between Toc159 and Toc34 gene family members.

Complex protein interactions can often be difficult to assess due to a variety of reasons. Limitations exist on purifying stable proteins, ensuring the recombinant proteins are similarly folded to *in vivo* conformations, as well as interaction with cofactors such as chaperones and nucleotide exchange factors do not always occur in recombinant systems. Often it is the case, as in the current study, that only a specific domain of a protein is used to assess the interaction potential which may pose limitations to the interactions observed if other domains possess regulatory or direct roles in interaction events. Specifically, the absence of such elements as the A- and M-domains of atToc159 and atToc33, the translocation channel Toc75, the membrane itself, as well as any transit peptides are potentially important elements for adequately characterising the interactions between atToc159 and atToc33. Toc159 and Toc34 gene families have been shown to interact but the ability to quantify the heterodimeric interaction

remains difficult possibly due to the absence of such components found only *in planta*. The development of techniques for testing these proteins in the presence of known interacting factors may also assist in illuminating the formation of Toc complexes.

#### **4.5 Integrating chloroplast protein translocation into the biological sciences**

Protein translocation into chloroplasts is a dynamic and complex process. Properly functioning protein translocons are important for the growth and development of all organisms at the cellular level. For example, the secretory pathway, protein import into mitochondria, as well as protein import into chloroplasts all are essential for the proper functioning and maintenance of cells and higher organism growth. More specifically, protein import into chloroplasts is essential for organelle identity. As discussed earlier, the protein complement of plastids enables the ability of plastids to interconvert between types. Thus the translocation machinery present in different plastid types is necessary for specific tissue development and function within plants. An understanding of how the different components of the import apparatus (e.g. Toc159 and Toc34) interact with one another should provide some insight into how the processes of protein import and plastid differentiation are linked. Proper development of plants is crucial as they are a vital component of life on earth, acting as a food source for countless species, and are capable of performing vital ecological processes such as photosynthesis. Thus, research invested in understanding the processes undertaken by plants is imperative. Gathering knowledge from a variety of biological disciplines the investigation regarding the development and function of the Tic/Toc system has established an ever-changing body of research. Utilizing tools from evolution (Bruce, 2000), plant and molecular biology (Bauer et al., 2000; Smith et al., 2002), as well as biochemistry and biophysics (Richardson, 2008) the investigation of protein translocation into chloroplasts reveals an integrative approach to biology. By gathering information from such sources, individual players within the translocation process have been predicted using protein function predicting software, confirmed using techniques such as immunoassays and pull-down assay, and characterized

using X-ray crystallography and plant knockout (*in vivo*) phenotypes. The integration of this information has led to the development of multiple models depicting the events of translocation. Although not one research group has been responsible for all variety of research, the integration of such information has allowed for significant advances in our understanding of protein translocation into chloroplasts. Implementing the knowledge of plant physiology has also become an increasingly dominant field of research. The advent of the biotechnology industry generated interest in understanding plant processes for the purpose of generating stronger, more efficient plants. The integration of knowledge and information from a variety of areas of biology and research interests, including protein import into chloroplasts, is no doubt a valuable asset to the generation of sustainable crops.

## 5. Conclusions

It has been demonstrated in the current study that the interaction between atToc159G and atToc33G can be seen as a stable heterodimeric structure when observed using BN-PAGE. The interaction between proteins has also been detected by observing changes to CD and FL spectra. The existence of homodimeric structures for both atToc159G and atToc33G have also been shown which, until recently were not considered in translocation models. Recently, a model has been proposed identifying the role of homodimers in translocation and current structural data gives no reason to reject such a hypothesis (Li et al., 2007). Observations identifying a higher order oligomer during atToc159G and atToc33G mixing has been hypothesized to reflect the interaction of atToc159G with multiple copies of atToc33G as suggested by stoichiometric studies of Toc complex composition. The generation of two single tryptophan mutants of atToc159G resulted in significant structural changes, which reveal the importance of W973F for indirectly maintaining the nucleotide binding domain and W1056F for its potential in regulating protein-protein interactions via the putative transit peptide binding site. Attempts to shift the monomer:dimer ratio using chelating agents EDTA and Chelex show the necessity of bound nucleotide to maintain a functional folded state. Irreversible changes to the structure of both atToc159G and atToc33G as a result of nucleotide loss were observed using CD and BN-PAGE revealing a significant shift towards monomer of up to 8-fold.

Further investigations should focus on utilizing and developing techniques which allow for the quantification of heterodimeric structures to assess the previous hypotheses about the formation of functionally distinct Toc complexes. Further investigations into the formation of homodimeric structures focused on elucidating their importance in Toc complex formation would also be of value.

## 6. References

- Abramoff, M.D., Magelhaes, P.J., and Ram, S.J. 2004. Image processing with ImageJ. *Biophotonics Intl.* **11**: 36-42.
- Agne, B., Infanger, S., Wang, F., Hofstetter, V., Rahim, G., Martin, M., Lee, D.W., Hwang, I., Schnell, D., and Kessler, F. 2009. A Toc159 import receptor mutant, defective in hydrolysis of GTP, supports preprotein import into chloroplasts. *J. Biol. Chem.* **284**: 8670-8679.
- Agne, B. and Kessler, F. 2009. Protein transport into organelles: The Toc complex way of protein import. *FEBS J.* **276**: 1156-1165.
- Aronsson, H., Boij, P., Patel, R., Wardle, A., Topel, M., and Jarvis, P. 2007. Toc64/OEP64 is not essential for the efficient import of proteins into chloroplasts in *Arabidopsis thaliana*. *Plant J.* **52**: 53-68.
- Baker, M.J., Frazier, A.E., Gulbis, J.M., and Ryan, M.T. 2007. Mitochondrial protein import machinery: Correlating structure with function. *TRENDS Cell Biol.* **17**: 456-464.
- Bauer, J., Chen, K., Hiltbrunner, A., Wehrli, E., Eugster, M., Schnell, D.J., and Kessler, F. 2000. The major protein import receptor of plastids is essential for chloroplast biogenesis. *Nature* **403**: 203-207.
- Becker, T., Jelic, M., Vojta, A., Radunz, A., Soll, J., and Schleiff, E. 2004. Preprotein recognition by the Toc complex. *EMBO J.* **23**: 520-530.
- Bradford, M. M. 1976. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **72**:248-254.
- Bruce, B.D. 2000. Chloroplast transit peptides: structure, function, and evolution. *TRENDS Cell Biol.* **10**: 440-447.
- Burstein, E.A., Vedenkina, N.S., Ivkova, M.N. 1973. Fluorescence and the location of tryptophan residues in protein molecules. *Photochem. & Photobiol.* **18**: 263-279
- Caliebe, A., Grimm, R., Kaiser, G., Lübeck, J., Soll, J., and Heins, L. 1997. The chloroplastic protein import machinery contains a Rieske-type iron-sulfur cluster and a mononuclear iron-binding protein. *EMBO J.* **16**: 7342-7350.

- Chen, D., and Schnell, D.J. 1997. Insertion of the 34-kDa chloroplast protein import component, IAP34, into the chloroplast outer membrane is dependent on its intrinsic GTP-binding capacity. *J. Biol. Chem.* **272**: 6614-6620.
- Chen, X., Smith, M.D., Fitzpatrick, L., and Schnell, D.J. 2002. In vivo analysis of the role of atTic20 in protein import into chloroplasts. *Plant Cell* **14**: 641-654.
- Chen, K and Li, H. 2007. Precursor binding to an 880-kDa Toc complex as an early step during active import of protein into chloroplasts. *Plant J.* **49**: 149-158.
- Chigri, F., Hormann F., Stamp A., Stammers D.K., Bolter B., Soll J., and Vothknecht U.C. 2006. Calcium regulation of chloroplast protein translocation is mediated by calmodulin binding to Tic32. *PNAS* **103**: 16051-16056.
- Chou, M.L., Chu C.C, Chen L.J., Akita M., and Li H.M. 2006. Stimulation of transit-peptide release and ATP hydrolysis by a cochaperone during protein import into chloroplasts. *J. Cell Biol.* **175**: 893-900.
- Chou, M.L., Fitzpatrick L.M., Tu S.L., Budziszewski G., Potter-Lewis S., Akita M., Levin J.Z., Keegstra K., and Li H.M. 2003. Tic40, a membrane-anchored co-chaperone homolog in the chloroplast protein translocon. *EMBO J.* **22**: 2970-2980.
- Clark S.A. and Theg S.M. 1997. A folded protein can be transported across the chloroplast envelope and thylakoid membranes. *Mol Biol Cell.* **8**: 923-934.
- Compton, L.A. and Johnson, Jr W.C. (1986) Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal. Biochem.* **155**, 155-167.
- Cui, X.L., Qin, C., and Zigler, J.S. 1994. Residual EDTA bound by lens crystallins accounts for their reported resistance to copper-catalyzed oxidative damage. *Arch. Biochem. Biophys.* **308**: 207-213.
- Dobberstein, B. and Blobel, G. 1975. Transfer of proteins across membranes I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes on murine myeloma. *J. Cell Biol.* **67**: 835-851.
- Dobberstein, B. and Blobel, G. 1975. Transfer of proteins across membranes II. Reconstitution of functional rough microsomes from heterologous components. *J. Cell. Biol.* **67**: 852-862.
- Duy, D., Wanner G., Meda A.R., von Wiren N., Soll J., and Philippar K. 2007. PIC1, an ancient permease in *Arabidopsis* chloroplasts mediates iron transport. *Plant Cell* **19**: 986-1006.

- Dyall, S.D., Brown M.T., and Johnson P.J. 2004. Ancient invasions: From endosymbionts to organelles. *Science* **304**: 253-257.
- Heins, L., Mehrle, A., Hemmler, R., Wagner, R., Kuchler, M., Hörmann, F., Sveshnikov, D. and Soll, J. 2002. The preprotein conducting channel at the inner envelope membrane of plastids. *EMBO J.* **21**: 2616-2625
- Hinnah, S.C., Hill K., Wagner R., Schlicher T., and Soll J. 1997. Reconstitution of a chloroplast protein import channel. *EMBO* **16**: 7351-7360.
- Hinnah, S.C., Wagner R., Sveshnikova N., Harrer R., and Soll J. 2002. The chloroplast protein import channel Toc75: pore properties and interaction with transit peptides. *Biophys. J.* **83**: 899-911.
- Hormann, F., Kuchler M., Sveshnikov D., Oppermann U., Li Y., and Soll J. 2004. Tic32, an essential component in chloroplast biogenesis. *J. Biol. Chem.* **279**: 34756–34762.
- Inaba, T., Li M., Alvarez-Huerta M., Kessler F., and Schnell D.J. 2003. atToc110 functions as a scaffold for coordinating the stromal events of protein import into chloroplasts. *J. Biol. Chem.* **278**: 38617-38627.
- Ivanova, Y., Smith M.D., Chen K., and Schnell D.J. 2004. Members of the Toc159 import receptor family represent distinct pathways for protein targeting to plastids. *Mol. Biol. Cell.* **15**: 3379-3392.
- 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.
- Kelly, S.M. & Price, N.C. 2000. The Use of Circular Dichroism in the Investigation of Protein Structure and Function. *Curr. Prot. & Peptide Sci.* **1**: 349-384.
- Kessler, F., Blobel G., Patel H.A., and Schnell D.J. 1994. Identification of two GTP-binding proteins in the chloroplast protein import machinery. *Science.* **266**: 1035-1039.
- Kessler, F. and Schnell, D.J. 2006. The function and diversity of plastid protein import pathways: A multilane GTPase highway into plastids. *Traffic* **7**: 248-257.
- Kiefer, F., Arnold K., Künzli M., Bordoli L., and Schwede T. 2009. The SWISS-MODEL repository and associated resources. *Nucleic acids Research* **37**: D387-D392.
- 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 Physiol.* **47**: 363-371.

- Kikuchi, S., Oishi M., Hirabayashi Y., Lee D.W., Hwang I., and Nakai M. 2009. A 1-megadalton translocation complex containing Tic20 and Tic21 mediates chloroplast protein import at the inner envelope membrane. *Plant Cell* **21**: 1781-1797.
- Koenig, P., Oreb M., Höfle A., Kaltofen S., Rippe K., Sinning I., Schleiff E., and Tews I. 2008. The GTPase cycle of the chloroplast import receptors Toc33/Toc34: Implications from monomeric and dimeric structures. *Structure*. **16**: 585-596.
- Koenig, P., Oreb M., Rippe K., Muhle-Goll C., Sinning I., Schleiff E., and Tews I. 2008. On the significance of Toc-GTPase homodimers. *J. Biol. Chem.* **283**: 23104-23112.
- Kouranov, A., Chen X., Fuks B., and Schnell D.J. 1998. Tic20 and Tic22 are new components of the protein import apparatus at the chloroplast inner envelope membrane. *J. Cell Biol.* **143**: 991-1002.
- Kovacheva, S., Bedard J., Patel R., Dudley P., Twell D., Rios G., Koncz C., and Jarvis P. 2005. *In vivo* studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. *Plant J.* **41**: 412-428.
- Kubis, S., Patel R., Combe J., Bédard 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. *Plant Cell.* **16**: 2059-2077.
- Küchler, M., Decker S., Hormann F., Soll J., Heins L. 2002. Protein import into chloroplasts involves redox-regulated proteins. *EMBO* **21**: 6136-6145.
- Kunkele, K., Heins S., Dembowski M., Nargang F.E., Benz R., Thieffry M., Walz J., Lill R., Nussberger S., and Neupert W. 1998. The preprotein translocation channel of the outer membrane of mitochondria. *Cell* **93**: 1009-1019.
- Lakowicz, J.R. 2006. Principles of Fluorescence Spectroscopy Third Edition. NY: Plenum Publishers. 529-573.
- Lee, J., Wang, F., and Schnell, D.J. 2009. Toc Receptor Dimerization Participates in the Initiation of Membrane Translocation during Protein Import into Chloroplasts. *J. Biol. Chem.* **284**: 31130-31141.
- Li, H.M. and L.J. Chen. 1997. A novel chloroplastic outer membrane-targeting signal that functions at both termini of passenger polypeptides. *J. Biol. Chem.* **272**: 10968-10974.
- Li H., Kesavulu, M.M., Su, P., Yeh, Y., and Hsiao, C. 2007. Toc GTPases. *J. Biomed. Sci.* **14**: 505-508.

- Martin, W., Stoebe B., Goremykin V., Hansmann S., Hasegawa M., and Kowallik K.V. 1998. Gene transfer to the nucleus and the evolution of chloroplasts. *Nature* **393**: 162-165.
- Millen, R.S., Olmstead R.G., Adams K.L, Palmer J.D., Lao N.T., Heggie L., Kavanagh T.A., Hibberd J.M., Gray J.C., Morden C.W., Calie P.J., Jermin L.S., and Wolfe K.H. 2001. Many parallel losses on *infA* from chloroplast DNA during angiosperm evolution with multiple independent transfers to the nucleus. *Plant Cell* **13**: 645-658.
- Miras, S., Salvi D., Piette L., Seigneurin-Berny D., Grunwald D., Reinbothe C., Joyard J., Reinbothe S., and Rolland N. 2007. Toc159- and Toc75-independent import of a transit sequence-less precursor into the inner envelope of chloroplasts. *J. Biol. Chem.* **282**: 29482-29492.
- Mokranjac, D. and Neupert W. 2008. Energetics of protein translocation into mitochondria. *Biochimica et Biophysica Acta* **1777**: 758-762.
- Nada, A., and Soll J. 2004. Inner envelope protein 32 is imported into chloroplasts by a novel pathway. *J. Cell. Sci.* **117**: 3975-3982.
- Nanjo, Y., Oka H., Ikarashi N., Kaneko K., Kitajima A., Mitsui T., Munoz F.J., Rodriguez-Lopez M., Boraja-Fernandez E., and Pozueta-Romero J. 2006. Rice plastidial N-glycosylated nucleotide pyrophosphatase/phosphodiesterase is transported from the ER-Golgi to the chloroplast through the secretory pathway. *Plant Cell.* **18**: 2582-2592.
- Palmer, J.D. 2003. The symbiotic birth and spread of plastids: how many times and whodunit? *J. Phycol.* **39**: 4-11.
- Parker, C.A. 1968. Photoluminescence of Solutions. *Elsevier*, New York.
- Petsko, G.A and Ringe D. 2006. Protein structure and function. *Sinauer Associates, Inc. Pub.* M.A., USA.
- Piston, D.W. and Kremers G. 2007. Fluorescent protein FRET: the good, the bad and the ugly. *TRENDS Biol. Sci.* **32**: 407-413.
- Provencher, S.W. and Gloeckner J. 1981. Estimation of globular protein secondary structure from circular dichroism. *Biochem.* **20**: 33-37.
- Qbadou, S., Becker T., Mirus O., Tews I., Soll J., and Schleiff E. 2006. The molecular chaperone Hsp90 delivers precursor proteins to the chloroplast import receptor Toc64. *EMBO* **25**: 1836-1847.

- Rahim, G., Bischof S., Kessler F., and Agne B. 2008. *In vivo* interaction between atToc33 and atToc159 GTP-binding domains demonstrated in a plant split-ubiquitin system. *J. Exp. Bot.* **60**: 257-267.
- Reddick, L.E., Vaughn M.D., Wright S.J., Campbell I.M., and Bruce B.D. 2007. *In vitro* comparative kinetic analysis of the chloroplast Toc GTPases. *J. Biol. Chem.* **282**: 11410-11426.
- Reumann, S., Inoue K., and Keegstra K. 2005. Evolution of the general protein import pathway of plastids. *Mol. Membrane Biol.* **22**: 73-86.
- Richardson, L. 2008. Structure-function analysis of the acidic domain of the Arabidopsis Toc159 receptors. *Master's Thesis*. University of Waterloo. Canada.
- Roberts, K. 2007. Handbook of plant science, volume 1. *Wiley-Interscience*, USA. Pp. 250-251.
- Ross, J.B.A, Szabo A.G., Hogue C.W.V. 1997. Enhancement of protein spectra with tryptophan analogs: fluorescence spectroscopy of protein-protein and protein-nucleic acid interactions. *Methods Enzym.* **278**: 151-190.
- Sato, S., Nakamura Y., Kaneko T., Asamizu E., and Tabata S. 1999. Complete structure of the chloroplast genome of *Arabidopsis thaliana*. *DNA Res.* **6**: 283-290.
- Schagger, H. and von Jagow G. 1991. Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* **199**: 223-231.
- Schleiff, E., Soll J., Kühlbrandt W., and Harrer R. 2003. Characterization of the translocon of the outer envelope of chloroplasts. *J. Cell Biol.* **160**: 541-551.
- Schleiff, E., Soll J., Sveshnikova N., Tien R., Wright S., Dabney-Smith C., Subramanian C., and Bruce B.D. 2002. Structural and guanosine triphosphate/diphosphate requirements for transit peptide recognition by the cytosolic domain of the chloroplast outer envelope receptor, Toc34. *Biochem.* **41**: 1934-1946.
- 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 Biol.* **7**: 303-304.
- Sharma, V.K. & Kalonia D.S. 2003. Steady-state tryptophan fluorescence spectroscopy study to probe tertiary structure of proteins in solid powders. *J. Pharma. Sci.* **92**: 890-899.
- Smith, M.D. 2006. Protein import into chloroplasts: an ever-evolving story. *Can. J. Bot.* **84**: 531-541.

- Smith, M.D., Hiltbrunner A., Kessler F., and Schnell D.J. 2002. The targeting of the atToc159 preprotein receptor to the chloroplast outer membrane is mediated by its GTPase domain and is regulated by GTP. *J. Cell Biol.* **159**: 833-843.
- Smith, M.D., Rounds C.M., Wang F., Chen K., Afithile M., Schnell D.J. 2004. AtToc159 is a selective transit peptide receptor for the import of nucleus-encoded chloroplast proteins. *J. Biol. Chem.* **165**: 323-334.
- Smith, M.D., and Schnell, D.J. 2004. Chloroplast Protein Targeting: Multiple Pathways for a Complex Organelle. In: *Protein Movement Across Membranes*. Ed. J. Eichler. 2004. Georgetown, TX: Landes Biosciences. 1-18.
- Sreerama, N. and Woody R.W. 2000. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**: 252-260.
- Stahl, T., Glockmann C., Soll J., and Heins L. 1999. Tic40, a new “old” subunit of the chloroplast protein import translocon. *J. Biol. Chem.* **274**: 37467-37472.
- Stengel, A., Benz P., Balsera M., Soll J., and Bolter B. 2008. TIC62 redox-regulated translocon composition and dynamics. *J. Biol. Chem.* **283**: 6656-6667.
- Stiller, J.W., Reel D.C., and Johnson J.C. 2003. A single origin of plastids revisited: convergent evolution in organellar genome content. *J. Phycol.* **39**: 95-105.
- Sun, Y., Forouhar F., Li H., , Tu S., Yeh Y., Kao S., Shr H., Chou C., Chen C., and Hsiao C. 2002. Crystal structure of pea Toc34, a novel GTPase of the chloroplast protein translocon. *Nature Struct. Biol.* **9**: 95-100.
- Sundaraaj, S., Guo A., Habibi-Nazhad B., Rouani M., Stothar P., Ellison M., and Wishart D.S. 2004. The CyberCell database (CCDB): a comprehensive, self-updating, relational database to coordinate and facilitate *in silico* modeling of *Escherichia coli*. *N. Acids. Res.* **32**: D293-D295.
- Sung, M., and Huh W. 2007. Bimolecular fluorescence complementation analysis system for *in vivo* detection of protein-protein interaction in *Saccharomyces cerevisiae*. *YEAST* **24**: 767-775.
- Szabo, A.G. 2000. Fluorescence principles and measurement. In: *Spectroscopy and spectrofluorimetry: a practical approach* 2<sup>nd</sup> Ed. *Oxford University Press*. N.Y., USA. Pp. 33-66.

- Teng, Y.S., Su Y.S., Chen L.J., Lee Y.J., Hwang I., and Li H.M. 2006. Tic21 is an essential translocon component for protein translocation across the chloroplast inner envelope membrane. *Plant Cell*. **18**: 2247-2257.
- Tsai, L.Y., Tu S.L., and Li H.M. 1999. Insertion of atToc34 into the chloroplastic outer membrane is assisted by at least two proteinaceous components in the import system. *J. Biol. Chem.* **274**: 18735-18740.
- Twine, S.M. and Szabo, A.G. 2003. Fluorescent amino acids analogs. *Methods in Enzymology*. **360**: 104-127.
- Villarejo, A., Buren S., Larsson S., Dejardin A., Monne M., Rudhe C., Karlsson J., Jansson S., Lerouge P., Rolland N., von Heijne G., Grebe M., Bako L., and Samuelsson G. 2005. Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. *Nat. Cell Biol.* **7**: 1224-1231.
- Vojta, A., Alavi M., Becker T., Hormann F., Kuchler M., Soll J., Thomson R., and Schleiff E. 2004. The protein translocon of the plastid envelopes. *J. Biol Chem.* **279**: 21401-21405.
- Weibel, P., Hiltbrunner, A., Brand, L., Kessler, F. 2003. Dimerization of Toc-GTPases at the chloroplast protein import machinery. *J. Biol. Chem.* **278**: 37321-37329.
- Whitmore, L. and Wallace B.A. 2004 DICHROWEB: an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research* **32**:W668-673.
- Wise, R.R. 2006. The Diversity of Plastid Form and Function. In: *The Structure and Function of Plastids*. Eds. R.R. Wise, and J.K. Hooper. Dordrecht: Springer. 3-26.
- Wise, R.R and Hooper J.K. 2007. The Structure and Function of Plastids. Volume 23. *Springer*. Netherlands. Pp. 75-98.
- Yeh, Y., Kesavulu, M.M., Li, H., Wu, S., Sun, Y., Konozy, E.H.E., Hsiao, C. 2007. Dimerization is Important for the GTPase Activity of Chloroplast Translocon Components atToc33 and psToc159. *J. Biol. Chem.* **282**: 13845-13853.
- Zhang, B., Zhang Y., Wang Z., and Zheng Y. 2000. The role of Mg<sup>2+</sup> cofactor in the guanine nucleotide exchange and GTP hydrolysis reactions of Rho family GTP-binding proteins. *J. Biol. Chem.* **275**: 25299-25307.

## Appendix 1: Primers for generating single tryptophan mutants of atToc159G using PCR

	Primer Set	Purpose of Primer Set		Sequence
atToc159G W973F	1	Generate the 5' fragment incorporating a 5' <i>NcoI</i> restriction site and the W973F mutation	S	GATACCATGGCTACATCTCAGGATG
			AS	TAGCGTTTTTAAATATGGATGTACC
	2	Generate the 3' fragment incorporating a 3' <i>XhoI</i> restriction site and the W973F mutation	S	GGTACATCCATATTTAAAAACGCTA
			AS	GTGGTGCTCGAGTCGGAAACCAAAT
atToc159G full length	3	Generate full length atToc159G mutant while maintaining the <i>NcoI</i> and <i>XhoI</i> restriction sites	S	GATACCATGGCTACATCTCAGGATG
			AS	GTGGTGCTCGAGTCGGAAACCAAAT
atToc159G W1056F	4	Generate the 5' fragment incorporating a 5' <i>NcoI</i> restriction site and the W1056F mutation	S	GATACCATGGCTACATCTCAGGATG
			AS	GCTGAGATCTAAAAGTTTGCCATT
	5	Generate the 3' fragment incorporating a 3' <i>XhoI</i> restriction site and the W1056F mutation	S	AATGGCCAAACTTTTAGATCTCAGC
			AS	GTGGTGCTCGAGTCGGAAACCAAAT

**Appendix 2: Calculation of monomer:dimer ratio for BN-PAGE using ImageJ.**

Wild-type proteins			Total Intensity	Proportion Dimer	Monomer : Dimer
<b>atToc159G</b>	Intensity of dimer	3679	12264	30.01	2 : 1
	Intensity of monomer	8584			
<b>atToc33G</b>	Intensity of dimer	2741	11408	24.03	3 : 1
	Intensity of monomer	8667			
<b>atToc159 with atToc33G</b>	Intensity of heterodimer	669	14479	4.62	N/A
	Intensity of remaining bands	4887			
atToc159G mutants			Total Intensity	Proportion Dimer	Monomer : Dimer
<b>atToc159G W973F</b>	Intensity of dimer	876	6433	13.62	7 : 1
	Intensity of monomer	5557			
<b>atToc33G W1056F</b>	Intensity of dimer	2974	12801	23.02	3 : 1
	Intensity of monomer	9854			
Chelex treated proteins			Total Intensity	Proportion Dimer	Monomer : Dimer
<b>atToc159G</b>	Intensity of dimer	609	5496	11.09	9 : 1
	Intensity of monomer	4887			
<b>atToc33G</b>	Intensity of dimer	554	13242	4.19	24 : 1
	Intensity of monomer	12687			