A plant split-ubiquitin system and TAP-tagging to study *in vivo* protein interactions in the chloroplast protein import machinery
A plant split-ubiquitin system and TAP-tagging to study

\textit{in vivo} protein interactions

in the chloroplast protein import machinery

Gwendoline RAHIM

UNIVERSITE DE NEUCHATEL

FACULTE DES SCIENCES

La Faculté des sciences de l'Université de Neuchâtel,
sur le rapport des membres du jury

MM. F. Kessler (directeur de thèse),
J.-M. Neuhaue, N. Rolland (Grenoble F)
et S. Baginsky (ETH Zürich)

autorise l'impression de la présente thèse.

Neuchâtel, le 2 avril 2009

Le doyen :
F. Kessler
Keywords: chloroplast, protein import, Toc complex, protein interactions, split-ubiquitin, TAP-tag

Mots-clefs: chloroplaste, import des protéines, complexe Toc, interactions protéine-protéine, split-ubiquitin, TAP-tag
<table>
<thead>
<tr>
<th>Chapter 2</th>
<th>Isolation of atToc33 and associated complexes from Arabidopsis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Presentation of the system</td>
</tr>
<tr>
<td>2</td>
<td>Generation of transgenic plants</td>
</tr>
<tr>
<td>3</td>
<td>Characteristics of plants expressing NTAPI-Toc33/Toc33G fusion proteins</td>
</tr>
<tr>
<td>3.1</td>
<td>Phenotype and chlorophyll measurements</td>
</tr>
<tr>
<td>3.2</td>
<td>Localisation of the fusion proteins</td>
</tr>
<tr>
<td>4</td>
<td>Purification of NTAPI-Toc33 and associated proteins</td>
</tr>
<tr>
<td>4.1</td>
<td>Small scale purification assay</td>
</tr>
<tr>
<td>4.2</td>
<td>Elution of NTAPI-Toc33 and NTAPI-Toc33G by TEV protease</td>
</tr>
<tr>
<td>5</td>
<td>Mass spectrometric protein identification</td>
</tr>
<tr>
<td>6</td>
<td>Discussion</td>
</tr>
<tr>
<td>6.1</td>
<td>Interaction of Toc33 with Toc159 and Toc75 depends on its membrane-association</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Differences in Toc proteins associated with NTAPI-Toc33 and NTAPI-Toc33G</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Complementation of the ppi1 phenotype and chlorophyll accumulation defect by NTAPI-Toc33</td>
</tr>
<tr>
<td>6.2</td>
<td>Hypothetic partners revealed by mass spectrometry</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Tic110 and chaperones</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Actin</td>
</tr>
<tr>
<td>6.2.3</td>
<td>α-tubulin and EF1α proteins</td>
</tr>
<tr>
<td>6.2.4</td>
<td>Calmodulin</td>
</tr>
<tr>
<td>6.2.5</td>
<td>Missing interaction partners</td>
</tr>
</tbody>
</table>

Chapter 2

**Isolation of atToc33 and associated complexes from Arabidopsis**

1 Presentation of the system
2 Generation of transgenic plants
3 Characteristics of plants expressing NTAPI-Toc33/Toc33G fusion proteins
   3.1 Phenotype and chlorophyll measurements
   3.2 Localisation of the fusion proteins
4 Purification of NTAPI-Toc33 and associated proteins
   4.1 Small scale purification assay
   4.2 Elution of NTAPI-Toc33 and NTAPI-Toc33G by TEV protease
5 Mass spectrometric protein identification
6 Discussion
   6.1 Interaction of Toc33 with Toc159 and Toc75 depends on its membrane-association
      6.1.1 Differences in Toc proteins associated with NTAPI-Toc33 and NTAPI-Toc33G
      6.1.2 Complementation of the ppi1 phenotype and chlorophyll accumulation defect by NTAPI-Toc33
6.2 Hypothetic partners revealed by mass spectrometry
   6.2.1 Tic110 and chaperones
   6.2.2 Actin
   6.2.3 α-tubulin and EF1α proteins
   6.2.4 Calmodulin
   6.2.5 Missing interaction partners
Materials and methods

1 Material

1.1 Plants

1.2 Bacteria and yeast strains

1.3 Chemicals

1.4 Oligonucleotides and Plasmids

1.5 Antibodies

2 Methods

2.1 Methods used for cloning

2.1.1 PCR, restriction digest and ligation

2.1.2 Transformation of chemically competent E. coli cells and plasmid isolation

2.2 DNA constructs

2.2.1 Yeast split ubiquitin cloning

2.2.2 Plant split-ubiquitin constructs

2.3 Stable transformation of Arabidopsis with TAP tag constructs

2.4 Plant growth

2.4.1 Growing plants in vitro

2.4.2 Growing Arabidopsis on soil

2.5 Yeast split-ubiquitin assay

2.6 Protoplasts transformation

2.7 Chlorophyll measurements

2.8 Plant genomic DNA extraction and PCR analysis

2.9 Plant proteins extraction and Western blot analysis

2.10 Separation of soluble and insoluble proteins

2.11 Immunolocalization on protoplasts

2.12 TAP tag purification

2.13 Protein identification by mass spectrometry
Abstract

Most chloroplast proteins are synthesized as precursor proteins in the cytosol. The import of these precursor proteins is mediated by molecular complexes located at the outer and inner membrane of the chloroplast. These complexes are called Toc (translocon at the outer envelope membrane) and Tic (translocon at the inner envelope membrane) respectively. In Arabidopsis, the Toc complex consists of three principle components: two homologous receptor GTPases, atToc159 and atToc33 and a protein-import channel: atToc75. During import, the two GTPases undergo complex interactions with precursor proteins and amongst themselves although precise mechanisms remain unknown.

In vitro studies revealed that Toc159 and Toc33 interact with each other via the dimerization of their GTP-binding domain (G-domain). Moreover, the crystal structure of the pea Toc33 ortholog, psToc34 indicates that it can stably homodimerize via its G-domain. However, neither Toc159/Toc33 heterodimers nor Toc33 homodimerization have been demonstrated in planta.

To get new insight into the in vivo interactions of Toc GTPases, we have developed a plant split-ubiquitin system. This method, originally developed for yeast, was adapted to study interactions between the Toc GTPases atToc159 and atToc33 in Arabidopsis and tobacco protoplasts. We also demonstrated that the peroxisomal membrane protein atPex11e, used initially as a model membrane protein in our system, self-interacts as does its yeast homolog. The plant split-ubiquitin system proves to be widely usable.

Another approach of this thesis was to get more information on the import mechanism via the identification of interaction partners of the Toc GTPase atToc33. atToc33 and proteins associated were isolated from Arabidopsis plants, using the tandem affinity purification (TAP) tag. We proved that this technique is suitable to purify Toc33, which encourages us to purify Toc proteins and complexes at a larger scale.
Résumé


Toc159 et Toc33 interviennent dans la reconnaissance du précurseur, mais les mécanismes exacts de reconnaissance et d’interactions de ces deux GTPases restent à découvrir. On sait cependant, grâce à des expériences menées in vitro, que Toc159 et Toc33 interagissent entre-elles par dimérisation par l’intermédiaire de leur domaine GTPasique (domaine G). De plus, la structure du cristal de psToc34, l’orthologue de Toc33 chez le pois, suggère que cette protéine a la possibilité de former un homodimère via son domaine G. Néanmoins, aucun homodimère de Toc33 ni aucune hétérodimerisation de Toc159 et Toc33 n’ont été encore observés in planta.

L’un des objectifs de mon travail était d’étudier ces interactions in vivo. Pour cela, nous avons développé chez Arabidopsis et le tabac, un système d’interaction protéine-protéine initialement crée dans la levure : le système "split-ubiquitin".

Afin d’en savoir plus sur les mécanismes d’import dans la plante, nous avons essayé dans un deuxième temps, d’identifier de nouvelles protéines participant à cet import. Pour cela, on a obtenu des plantes exprimant la protéine Toc33 fusionnée à une étiquette TAP-(Tandem affinity purification) afin d’isoler, purifier puis identifier les protéines qui interagissent avec cette Toc GTPase.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-protein</td>
<td>antibodies recognizing the specified protein</td>
</tr>
<tr>
<td>A-domain</td>
<td>acidic domain</td>
</tr>
<tr>
<td>atToc159G</td>
<td>GTP binding domain of atToc159</td>
</tr>
<tr>
<td>atToc33G</td>
<td>GTP binding domain of atToc33</td>
</tr>
<tr>
<td>atToc159A</td>
<td>acidic-domain of atToc159</td>
</tr>
<tr>
<td>at, <em>A. thaliana,</em></td>
<td><em>Arabidopsis thaliana</em></td>
</tr>
<tr>
<td>bp</td>
<td>base pairs</td>
</tr>
<tr>
<td>CaMV</td>
<td>cauliflower mosaic virus</td>
</tr>
<tr>
<td>CBP</td>
<td>Calmodulin binding protein</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>Col</td>
<td>Columbia</td>
</tr>
<tr>
<td>Cub</td>
<td>C-terminal half of ubiquitin</td>
</tr>
<tr>
<td>DNA</td>
<td>desoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>desoxynucleotide triphosphate</td>
</tr>
<tr>
<td>DUBs</td>
<td>deubiquitinating enzymes</td>
</tr>
<tr>
<td><em>E. coli</em></td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EGTA</td>
<td>ethyleneglycoltetraacetic acid</td>
</tr>
<tr>
<td>FITC</td>
<td>fluorescein isothiocyanate</td>
</tr>
<tr>
<td>G-domain</td>
<td>GTP-binding domain</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>HA</td>
<td>hemagglutinin</td>
</tr>
<tr>
<td>His6</td>
<td>hexahistidinyl-tag</td>
</tr>
<tr>
<td>IgG</td>
<td>immunoglobulin G</td>
</tr>
<tr>
<td>kb</td>
<td>kilo base pairs</td>
</tr>
<tr>
<td>kDa</td>
<td>kilo Dalton</td>
</tr>
<tr>
<td>LB</td>
<td>Luria Bertani</td>
</tr>
<tr>
<td>M-domain</td>
<td>Membrane domain</td>
</tr>
<tr>
<td>MS</td>
<td>Murashige and Skoog</td>
</tr>
<tr>
<td>Nub</td>
<td>N-terminal half of ubiquitin</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate-buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>Pex</td>
<td>Peroxisomal protein</td>
</tr>
<tr>
<td>ppi</td>
<td>plastid protein import</td>
</tr>
<tr>
<td>protA</td>
<td>protein A</td>
</tr>
<tr>
<td>PRK</td>
<td>phosphoribulokinase</td>
</tr>
</tbody>
</table>
ps  
_Pisum sativum_

rt  
room temperature

RubisCO  
ribulose-1,5-bisphosphate carboxylase oxygenase

_S. cerevisiae_  
_Saccharomyces cerevisiae_

SDS  
sodium dodecyl sulfate

SDS-PAGE  
SDS-polyacrylamide gel electrophoresis

SRP  
signal recognition particle

TAP tag  
tandem affinity purification tag

TBS  
tris-buffered saline

T-DNA  
transfer DNA

TEV  
_Tobacco Etch Virus_

Tic  
translocon at the inner chloroplast membrane

TM  
transmembrane helix

Toc  
translocon at the outer chloroplast membrane

Tris  
tris (hydroxymethyl) aminomethane

TRITC  
tetramethylrhodamine isothiocyanate

T0 lines  
Arabidopsis plants that were transformed by floral dip

T1 lines  
first generation of seeds after plant transformation by floral dip

U  
unit

ub  
ubiquitin

UBPs  
ubiquitin specific proteases

UV  
ultra violet

v/v  
volume per volume

wt  
wild type

w/v  
weight per volume

atToc…, atTic…  
indicates the Toc-, Tic-protein of _A. thaliana_

psToc…, psTic…  
designates the Toc-, Tic-protein of _P. sativum_

Toc…, Tic…  
refers to Toc-and Tic-proteins of both _A. thaliana_ and _P. sativum_ unless specified otherwise

upper case, italic  
gene, wild type allele (e.g. _TOC33_)

lower case, italic  
mutant allele(s) (e.g. _ppi1_)

regular case  
protein (e.g. atToc33)
Introduction

1 Chloroplasts, the green plastids

1.1 Characteristics and distribution of plastids

Plastids are specialised organelles hosted by plant cells. They provide carbon and energy to the plant as a result of photosynthetic assimilation (Pyke, 1999). Biosynthesis events such as the synthesis of lipids, nucleic acids and amino acids take place in plastids (Galili, 1995, Ohlrogge and Browse, 1995). Originally, plastids were categorized in three groups by A.F.W Schimper in 1885 (Mühlethaler, 1971) according to their colour: leucoplasts (white plastids), chromoplasts (yellow plastids) and chloroplasts (green plastids). Chloroplasts are located in the plant green tissues as they contain chlorophyll pigments and are the site of photosynthesis. Chloroplasts are also present during the first days of embryo development (Mansfield and Briarty, 1992). Leucoplasts store starch (amyloplasts), proteins (proteinoplasts) or lipids (elaioplasts) and are present in roots and non photosynthetic tissues but also in petals (Pyke and Page, 1998). Chromoplasts are red-, orange, and yellow-coloured plastids which contain carotenoid pigments. They are located in fruits and flowers (Marano, et al., 1993).

All these different groups of plastids derived from proplastids, which are present in meristematic cells (Leon et al., 1998). One meristematic cell is proposed to contain 10 to 20 proplastids (as reviewed by Pyke, 1999). In the presence of light, proplastids synthesise chlorophyll and become chloroplasts. Alternatively, if seedlings are grown in the dark, proplastids are differentiated to etioplasts.

1.2 Structure of the chloroplast

Chloroplasts are about 5-10μm in diameter and 3-4μm in thickness, present in each cell at numbers of 10 to more than 100 depending on the plant species (as reviewed by Lopez-Juez and Pyke, 2005). They are bounded by two envelope membranes (Fig.1). The photosynthetic apparatus is localised in an internal system, the thylakoids, also enclosed by a membrane. Thylakoids usually forms sacks which are arranged in stacks, called grana, but individual thylakoids are also found (stromal thylakoids). They are embedded in the stroma, the chloroplast aqueous matrix, together with DNA, ribosomes, starch and lipid droplets called plastoglobuli.
Chloroplasts have two lipid membranes separated by an intermembrane space that enclose the stroma. The stroma contains the thylakoids, arranged in stacks called grana.

**Fig.1: The chloroplast**

A: Transmission electron microscopy of the chloroplast (University of Wisconsin Botany website http://botit.botany.wisc.edu/)

B: Schematic representation of the chloroplast (Campbell, N.A, 1996)
The proteins controlling photosynthesis (photosystems I and II, the cytochrome b$_6$/f complex and the ATP synthase) are positioned in the thylakoid membranes. The grana disposition of thylakoids is delimiting the two photosystems: the photosystem I is in contact with the stroma whereas the photosystem II is limited to the granal membranes only.

Chloroplast and thylakoid envelope membranes are rich in galactolipids, whereas the other plant membranes are mainly composed of phospholipids (Jarvis et al., 2000). The chloroplast outer envelope is permeable to molecules up to 10kDa while the inner envelope is more selective (Lopez-Juez and Pyke, 2005).

2 Evolution

2.1 “Little workers, green slaves”

The hypothesis that plastids have an endosymbiont origin already arose in 1905, when Mereschkowsky suggested that plastids have cyanobacterial origins and act as “little workers, green slaves” within the cell (Martin and Kowallik, 1999). It is now commonly accepted that plastids emerged more than 1 billion years ago from one symbiogenetic event when a single ancestral cyanobacterium (Martin and Hermann, 1999; Palmer and Delwiche, 1998) integrated into the ancestral host cell (Cavallier-Smith, 2000).

This was confirmed by studies and comparison of genes coding for chloroplast proteins with genes that have a cyanobacterial origin (Sugita et al., 1997). Proteins sequence comparisons also demonstrate that plants and red algae are phylogenetically related (Moreira et al., 2000; McFadden and van Dooren, 2004).

Due to their endosymbiont origins, plastid division is similar to bacterial fission. Plastids replicate from pre-existing plastids and plastid division is the result of a series of events which include plastid division components of both plant and cyanobacterial origin (Pyke, 1999).

2.2 Genome evolution

The plant genome was also affected by this endosymbiotic event. Indeed, analysis of the nuclear genome of Arabidopsis showed that about 18% of the protein-coding genes have a cyanobacterial origin (Martin et al., 2002). Compared to cyanobacteria that has approximately 3000 genes (Kaneko et al., 1996), chloroplast genomes contain between 120 and 135 genes which encode less than 80 proteins, mainly components of the
photosynthesis protein complexes (Lopez-Juez and Pyke, 2005). Some of the genes that were not essential anymore for the endosymbiont disappeared during evolution and the other genes of the endosymbiont were transferred to the plant nuclear genome (Martin et al., 1998, Timmis et al., 2004).

The majority of the proteins required for plastid functions is thus encoded in the nucleus and translated on cytoplasmic ribosomes. These nuclear encoded proteins need to be targeted to the chloroplast surface. Proteins destined to be imported have to pass through the two envelope membranes before carrying on their journey inside the chloroplast.

3 Chloroplast protein import machinery

3.1 Targeting of the preproteins

The majority of the imported proteins are synthesised as precursor proteins (preproteins) with a cleavable N-terminal transit peptide. This peptide is essential and sufficient to address the preprotein to the plastid outer membrane and to transfer it across the chloroplast translocation machinery (Becker et al., 2005). Proteins targeted to the thylakoid membrane or lumen bear bipartite N-terminal presequences, with two domains, one for each compartment the protein will be targeted to (Robinson et al., 2001).

However, most of the outer envelope membrane proteins are not synthesised with a transit peptide. This is also true for some chloroplast-imported proteins (Kleffmann et al., 2004).

3.2 The chloroplast transit peptides

Compared to other transit peptides that belong, for example, to the secretory pathway, or mitochondrial proteins, chloroplasts targeting signals lack consensus features. They are made up of 20 to 80 amino acids, contain a low amount of acidic and hydrophobic residues, but many of them are hydroxylated and positively charged. Transit peptides are rich in serine and threonine that may act as phosphorylation sites. The presence of such sites raised the hypothesis that transit peptides could interact with cytosolic and chloroplastic chaperones required for targeting and import (von Heijne and Nishikawa, 1991). Indeed more than 75% of transit peptides are predicted to have at least one binding site for DnaK, the 70 kDa heat shock protein (Hsp70) from E. coli. In vitro
protein binding experiments detected the interaction between DnaK and the transit peptide of two chloroplastic imported proteins: ferredoxin NADP oxidoreductase and the small subunit of ribulose bisphosphate carboxylase/oxygenase (RuBisCO) (Ivey and Bruce, 2000; Rial et al., 2000). Hsp70s are ATPase proteins that play a key role in protein folding and may bind to the transit peptides of preproteins to prevent aggregation and give them the spatial conformation (unfolded state) necessary for import through the chloroplast membrane. However altering the binding capacity of precursor proteins to the Hsp70 chaperone does not affect import (Rial et al., 2003). This indicates that the exact role of chaperones during chloroplast protein translocation remains to be solved.

3.3 Guidance complex

Transit peptides are rich in the hydroxylated amino acids serine and threonine. These residues are known to be a site for phosphorylation and are also binding sites for 14-3-3 proteins that belong to a family of highly conserved proteins of 30 kDa expressed in all eukaryotic cells. 14-3-3 proteins play a role in cell signalling pathways and are proposed to be phospho-dependent chaperones (as reviewed by Yaffe, 2002 and by Ferl, 2004). It has been proposed that precursors would bind to 14-3-3 proteins after phosphorylation of their transit peptide, together with Hsp70 and maybe unidentified components to form a guidance complex (May and Soll, 2000). However, the role of the guidance complex in preprotein targeting remains hypothetical, as the mutation of the phosphorylation sites did not impair preprotein import (Nakrieko et al., 2004).

3.4 The existence of a cytosolic receptor

An alternative to the guidance complex involves a cytosolic form of Toc159, one of the preprotein receptors (that form the Toc complex) located at the outer membrane of the chloroplast. This complex and its components will be described later.

It has been suggested that Toc159 exists in a cytosolic, soluble form as well as in a chloroplast membrane bound form (Hiltbrunner et al., 2001b). The soluble form of Toc159 would then recognise and bind to the preprotein and “guide” it to the chloroplast membrane to be translocated. Recent findings have indeed shown that soluble Toc159 directly interacts with preproteins (Smith et al., 2004). In vivo data are still missing to confirm this hypothesis and to identify the components that are involved in targeting of the soluble Toc159-preprotein complex to the membrane.
Fig. 2: Targeting of chloroplast precursor proteins (from Becker *et al.*, 2005)

Four different pathways are proposed for targeted proteins that have no transit peptide (A), a non-phosphorylated transit peptide (B), a phosphorylated transit peptide (C), or are guided by the soluble form of Toc159 (D).
3.5 Four different ways of preprotein targeting

To conclude, there are four different hypotheses concerning the targeting of preproteins to the chloroplast membrane (as reviewed by Becker et al., 2005 and in Fig 2). Chloroplastic proteins without transit peptides are directly inserted into the membrane. Non-phosphorylated preproteins may be imported via cytosolic compounds that remain to be identified, whereas phosphorylated ones may bind to the guidance complex. The soluble form of the outer envelope receptor Toc159 is also proposed to be one of the cytosolic components necessary for the addressing of precursors.

4 The Chloroplast translocation machinery

The import of preproteins is facilitated by translocon complexes located in both chloroplast envelope membranes. These complexes are called Toc (translocon at the outer envelope membrane) and Tic (translocon at the inner envelope membrane) respectively. The numbers of the different Toc and Tic components correspond to their molecular mass in kDa.

4.1 The Toc complex

In pea, the initial plant model that was used to study chloroplast protein import, the Toc complex consists of Toc159, Toc34, Toc75, Toc12 and possibly Toc64 (Fig.3), but only the three first proteins are considered as the core components of the Toc complex as they were sufficient to allow protein translocation using artificial lipid vesicles (Schleiff et al., 2003a). The core complex is proposed to have a molecular weight of 500kDa and a stoichiometry of 1/4/4 between Toc159/Toc75/Toc34 (Schleiff et al., 2003b). These three proteins that form this hetero-oligomeric complex in pea will be described in more detail below.

Toc64 was characterised in pea by co-purification with Toc159, Toc34 and Toc75 (Sohrt and Soll, 2000), however in a later work performed by the same group, Toc64 did not co-purify with the other Toc components suggesting that the interaction between Toc64 and the Toc core component is transient (Schleiff et al., 2003b). Toc64 is made of three tetratricopeptide repeat (TPR) motifs that are known to mediate protein-protein interactions and interaction between proteins and chaperones (reviewed by Bedard and Jarvis, 2005). The role of Toc64 remains elusive, but because of its TPR motifs, it could
Fig. 3. Components of the chloroplast protein import machinery (adapted from Kessler and Schnell, 2006, and Lopez-Juez and Pyke, 2005).

Targeting of preproteins to the chloroplast surface is still not well known, but could involve a guidance complex composed of 14-3-3 proteins and a chaperone Hsp70 as well as Toc64. The Toc core components consist of two GTPase receptors, Toc159, Toc34 and a channel protein Toc75. Preproteins are translocated across the Toc complex, and then bind to another chaperone, Hsp70, and Tic22 in the intermembrane space, to guide them towards the Tic complex. Tic110 and (or) Tic20 form the Tic channel. The stromal part of Tic110 is proposed to bind preproteins and recruit chaperones Hsp93 and Cpn60 with the help of Tic40. Tic55, Tic62 and Tic 32 may regulate the protein import via a redox cycle.

Once in the stroma, the transit peptide is cleaved by a stromal processing peptidase (SPP). Import into the thylakoids involves different pathways depending on the nature of the protein imported. Once in the thylakoid lumen, the “thylakoidal” signal peptide is cleaved by a thylakoid processing peptidase (TPP).
play a role in the guidance complex as a docking site for preproteins (Qbadou et al., 2006; Qbadou et al., 2007; Schlegel et al., 2007; Sohrt and Soll, 2000). However, in recent studies in the moss Physcomitrella patens, Toc64 is proposed to have no role in protein import but to be involved in chloroplast morphology, as mutants were not impaired in import but had a difference in plastid shape (Hofmann and Theg, 2005). In Arabidopsis, efficiency of import is also not affected by Toc64 mutations (Aronsson et al., 2007).

Once the precursor proteins reach the chloroplast outer membrane, their recognition and binding are performed by Toc159 and Toc34, two homologous GTPases. Hydrolysis of GTP occurs during import but the precise mechanisms of the two GTPases (Toc159 and Toc34) are still unknown (Kessler and Schnell, 2004). Once recognized by the Toc complex, preproteins are translocated across the outer membrane passing through Toc75, the protein import channel.

ATP and GTP are required for the irreversible binding of preprotein to the translocon complex (Olsen et al., 1989; Young et al., 1999) and ATP is necessary for its translocation (Theg et al., 1989).

In the intermembrane space, Toc12 has been proposed to recruit Hsp70, which will increase the chaperone affinity for the preprotein that will be addressed to the Tic complex with the help of Tic22 (Becker et al., 2004a).

Toc12 is anchored to the outer envelope membrane, but its C-terminal part is located in the intermembrane space. Toc12 contains a J-domain and stimulates the ATPase activity of DnaK in vitro (Becker et al., 2004a).

4.2 The Tic complex

Toc and Tic proteins are believed to interact and associate with each other. It has been suggested that these Toc-Tic super-complexes are formed during preprotein import across the chloroplast membrane (Kessler and Schnell, 2006).

The Tic complex components are: Tic110, Tic62, Tic55, Tic40, Tic32, Tic22 and Tic20 (as reviewed by Keegstra and Cline, 1999) (Fig.3).

Tic110 is proposed to be the inner envelope membrane protein import-channel (Heins et al., 2002) together with Tic20 (Kouranov et al, 1998). Recently, Tic21, a protein with topological similarity to Tic20, was identified as a third putative channel component (Teng et al., 2006), although the role of this component in chloroplast protein import has been disputed (Duy et al., 2007). The stromal part of Tic110 is supposed to bind preproteins as this region of Tic110 contains a transit peptide binding site (Inaba et al., 2003). While binding to preprotein, Tic110 would recruit chaperones Hsp93 and Cpn60 with the help of Tic40 (Kessler and Schnell, 2006). These chaperones probably keep the preprotein in a correct folding conformation and give the driving force to complete the
translocation into the stroma. The role of Tic22 is still not known but as it is facing the inter membrane space, this protein could be a preprotein receptor or it could help formation of Toc-Tic super-complexes during translocation (Kouranov et al., 1998).

Tic62, Tic55 and Tic32 are believed to be redox components that are proposed to serve as regulators of the translocation reaction (Caliebe et al., 1997; Kuchler et al., 2002; Hormann et al., 2004; Bédard and Jarvis, 2005; Jarvis, 2008). Indeed, Tic62 interacts with a ferredoxin-NAD(P)\(^+\) oxidoreductase that can be regulated by photosynthetic electron transfer (Küchler et al., 2002) and Tic55 is predicted to contain a Rieske-type iron-sulfur cluster that catalyses electron transfer reactions (Caliebe et al., 1997). The role of Tic32 is still unclear but this protein is essential as Arabidopsis mutant plants that do not express this protein are embryo lethal (Hormann et al., 2004).

4.3 Import towards the thylakoids

Once in the stroma, a stromal processing peptidase (SPP) cleaves the transit peptide. The preprotein becomes a mature protein unless its fate is to be targeted to the thylakoid.

Different pathways are involved in the targeting of thylakoid preproteins (as reviewed by Robinson et al., 2001). Thylakoid membrane proteins either integrate via the signal recognition particle pathway (SRP), or, integrate spontaneously without the assistance of any known apparatus. Proteins addressed to the lumen are imported by the Sec pathway or the Twin-arginine translocation (Tat) pathway. The Sec pathway is comparable to the way prokaryotes export or secrete proteins. The Sec pathway involves a stromal SecA, ATP hydrolysis and a Sec complex that binds to the membrane. The Tat pathway is also called the ΔpH-dependent pathway, as the pH gradient across the thylakoid drives the translocation. It translocates fully folded proteins into the thylakoids; ATP, and maybe GTP, are involved in this import. When the proteins are in the lumen, the thylakoid targeting signal peptide is cleaved by a thylakoid processing peptidase (TPP).

4.4 Chloroplast protein import can be divided into three stages

The first step of import into chloroplast does not require energy and consists of reversible precursor protein association with the chloroplast outer membrane (Perry and Keegstra, 1994; Ma et al., 1996).

The second step consists of the binding of the precursor protein to the chloroplast import machinery to form an “early import intermediates”. This binding is irreversible
and requires low concentrations of either hydrolyzable GTP or ATP (100µM) (Olsen et al., 1989; Olsen and Keegstra, 1992; Schnell and Blobel, 1993). The required ATP is hydrolysed in the intermembrane space (Olsen and Keegstra, 1992).

The ultimate stage is the complete translocation of the precursor protein through both outer and inner envelope membrane and, when arrested, is called late translocation intermediate. This step requires concentrations of ATP of 1 to 3mM (Pain et Blobel, 1987; Theg et al., 1989), which is probably consumed by ATP hydrolyzing chaperones such as Hsp60, Hsp70 or Hsp100 (Kessler and Blobel 1996).

### 4.5 Import of proteins independent of the translocon complexes

Further pathways lead to import of chloroplast proteins. The majority of outer membrane proteins that do not possess a transit peptide insert spontaneously, or, by interaction with lipids of the outer envelope membrane (Keegstra and Cline, 1999).

Some proteins transported to the inner envelope membrane also lack a transit peptide. For example, the targeting information of Tic32 (Nada and Soll, 2004) and the chloroplast envelope quinone oxidoreductase protein (Miras et al., 2007), is present within the mature protein. These two proteins are not targeted through the Toc complex apparatus (Nada and Soll, 2004; Miras et al., 2007).

Interestingly, the secretory pathway also plays a role in chloroplast import as some proteins are N-glycosylated before being imported (Villajero et al., 2005), which is not so surprising as some chloroplast proteins were predicted to have a signal peptide for ER targeting (Kleffmann et al., 2004).

A substrate dependent import also exists. This is the case for one of the light-dependent enzymes that control chlorophyll synthesis: NADPH-protochlorophyllide (PChlide) oxidoreductase A (PORA). PORA import is suggested to be dependent of its substrat PChlide (Reinbothe et al. 1995) and to involve atToc33, the Arabidopsis homologue of psToc34 (Reinbothe et al. 2005). However, some studies have shown that the import is substrate-independent (Aronsson et al., 2000) and may involve the guidance complex (Schemenewitz et al., 2007). Thus, it has been proposed that both pathways are used depending on the plant development stage (Schemenewitz et al., 2007).
5 The different components of the Toc core complex

5.1 Toc75

Although Toc75 is located at the outer membrane of the chloroplast, it is synthesised with a large bipartite transit peptide (Tranel and Keegstra, 1996). The N-terminal part targets the protein and is cleaved by a stromal processing peptidase; the C-terminal part of the transit peptide prevents targeting to the stroma and mediates envelope targeting (Tranel and Keegstra, 1996; Inoue and Keegstra, 2003).

Toc75 is one of the protein import components (Perry and Keegstra, 1994; Schnell et al., 1994; Tranel et al., 1995) that is predicted to form a β-barrel structure and to act as a cation selective conducting channel (Hinnah et al., 1997; Hinnah et al., 2002).

In the plant model Arabidopsis thaliana, three genes encode proteins homologous to pea Toc75 (psToc75): atTOC75-I, atTOC75-III, and atTOC75-IV but only atTOC75-III seems to code for the ortholog of psToc75 (Jackson-Constant and Keegstra, 2001).

The discovery of syn75, an homolog of pea Toc75 in the cyanobacterium Synechocystis PCC6803 (Bölter et al, 1998; Reumann et al., 1999) support the fact that Toc75 has a prokaryotic origin, on the opposite, no homolog of Toc159 and Toc34 are found in Synechocystis, consequently these two GTPases are believed to have an eukaryotic origin (Reumann and Keegstra, 1999; Dyall et al., 2004).

5.2 Toc GTPases

The common feature of the two GTPases of the Toc complex, Toc159 and Toc34, is a conserved GTP binding domain (G-domain). These GTPases of eukaryotic origin belong to a new family of GTPases (Sun et al., 2002, Aronsson et al., 2003) that also belong to the paraseptin family (as reviewed by Weirich et al., 2008).

Toc159 is composed of an N-terminal acidic domain (A-domain), followed by a central G-domain and by a C-terminal domain (M-domain) that is anchored to the membrane. Toc34 is constituted almost exclusively of its G-domain, and is anchored in the outer membrane by its C-terminal part which is composed of a stretch of hydrophobic amino acids.

In Arabidopsis, two genes encode proteins homologous to pea Toc34 (Fig.4): atTOC33 and atTOC34 (Jarvis et al., 1998; Hiltbrunner et al., 2001a, Jackson-Constan and Keegstra, 2001), four genes encode homologs of psToc159: atTOC159, atTOC132, atTOC120 and atTOC90 (Bauer et al., 2000; Hiltbrunner et al., 2001a; Hiltbrunner et
The members of the family of GTP binding proteins of Arabidopsis fall into two subgroups, the homologues of pea Toc159 (atToc159, atToc132, atToc120 and atToc90) and the homologues of pea Toc34 (atToc34 and atToc33). They all share a GTP binding-domain (G-domain). Toc159 family proteins have an additional N-terminal acidic domain (A-domain) with the exception of Toc90 and a C-terminal membrane domain (M-domain), whereas the Toc34 family has a C-terminal transmembrane helix (TM). The numbers indicate amino acids.

**Fig.4: GTP binding proteins of the Toc-complex in A. thaliana:** (from Hiltbrunner et al., 2001a and Weibel, 2005)
al., 2004). atToc159 and atToc33 are the orthologs of psToc159 and psToc34 respectively. All the homologs of psToc159, apart from atToc90, share an N-terminal A-domain.

Arabidopsis is not the only plant where homologs of pea Toc159 and Toc34 proteins are present. Analysis of genomic databases identified two or three Toc34 homologs in at least seven plant species such as spinach, rape seed, seed, potato, tomato, poplar tree, maize, and in the moss Physcomitrella patens, moreover, homologs of Toc159, Toc132 and Toc120 are known in rice and Physcomitrella patens (Voigt et al., 2005).

5.2.1 Toc159 and its homologs

Due to its proteolytic sensibility, Toc159 was first identified as an 86kDa protein (Waegemann and Soll, 1991) and was then initially named Toc86 (Hirsch et al., 1994; Kessler et al., 1994). Toc159 is believed to be the primary receptor of preproteins (Hirsch et al., 1994, Perry and Keegstra, 1994), and, as discussed earlier, binding to preproteins may happen when Toc159 is in a soluble cytosolic form.

The function of the A-domain of Toc159 is still not known, but in Arabidopsis, proteins lacking this domain can complement the atToc159 knock-out mutant, ppi2 (Lee et al., 2003). In Arabidopsis protoplasts, import is not affected if proteins are lacking this domain (Lee et al., 2003) and Toc159 is still located at the chloroplast outer membrane, even when this domain is missing (Bauer et al., 2002).

The M-domain anchors the protein to the chloroplast outer membrane, and is sufficient to complement the defect in protein import in Arabidopsis protoplasts (Lee et al., 2003). This complementation is not complete though, as complemented plants have 50% less chlorophyll than wild type plants (Lee et al., 2003).

The G-domain contains GTP binding motifs as well as a dimerisation motif (Sun et al., 2002). The G-domain binds GTP and is involved in the targeting of Toc159 to the membrane in the presence of GTP (Bauer et al., 2002; Smith et al., 2002, Lee et al., 2003). Indeed, only this domain is sufficient to target the protein to the chloroplast outer membrane in Arabidopsis protoplasts (Bauer et al., 2002). Targeting and insertion of Toc159 needs GTP binding and hydrolysis (Bauer et al., 2002; Smith et al., 2002).

It has been proposed that both the G- and the M-domain play a role in Toc159 targeting to the chloroplast membrane (Lee et al., 2003).

In Arabidopsis, the atToc159 knock-out mutant, ppi2 (ppi stands for plastid protein import), has an albino phenotype, plastids do not develop and the plants cannot grow autotrophically, moreover, this mutation is seedling lethal (Bauer et al., 2000). This mutant is the result of a T-DNA insertion that disrupts the atTOC159 gene. The other members of the import apparatus are expressed and functional in this mutant; it appears
that Toc159 is essential for chloroplast biogenesis, especially as accumulation of photosynthetic proteins is dramatically reduced. However, import of non-photosynthetic proteins does not seem to be affected in ppi2 plants. It has been suggested that the three other homologs of psToc159, atToc132, atToc120 and atToc90 partially compensate for the absence of atToc159, or that these proteins are dedicated to the import of non-photosynthetic proteins (Bauer et al., 2000). Knock-out mutants of either atToc132, atToc120 or atToc90 show no specific phenotype, this suggests that their function is not essential and that they are functionally redundant (Hiltbrunner et al., 2004; Ivanova et al., 2004; Kubis et al., 2004). Preprotein binding studies confirmed that atToc132 and atToc120 have specificity for non-photosynthetic proteins (Ivanova et al., 2004).

The double mutant atToc132/atToc120 is lethal (Ivanova et al., 2004), or has a phenotype similar to ppi2 (Kubis et al., 2004), which confirms that these two proteins are redundant. atToc159 cannot complement the atToc132/atToc120 mutant phenotype and the double mutant atToc159/atToc132 is embryo lethal (Kubis et al., 2004); these findings support the fact that the members of the Toc159 family in Arabidopsis are involved in different import pathways.

Double knock out atToc90/atToc132, atToc90/atToc120 have no phenotype which is surprising as the atTOC90 gene is expressed with high levels within the plant (Kubis et al., 2004). Therefore, the role of atToc90 still has to be defined, but atToc90 definitely does not seem to be essential for import and maybe even is not acting in chloroplast protein import (Hiltbrunner et al., 2004). These results also show that there is no functional redundancy between atToc90/atToc132 and atToc90/atToc120.

5.2.2 Toc34

Toc34 is composed of one C-terminal transmembrane helix and of a G-domain similar to the G-domain of Toc159. Toc34 is synthesized with no cleavable presequence and inserts directly into the lipid layer of the chloroplast outer envelope membrane with the help of GTP (Qbadou et al., 2003). Crystal structure of the G-domain of pea Toc34 and in vitro studies, suggests homodimerisation of Toc34 in a GDP-bound form but also heterodimerisation with the G-domain of Toc159 (Sun et al., 2002; Weibel et al., 2003). Integration and assembly of Toc159 in the Toc complex is proposed to be mediated by Toc34 GTPase activity with the help of Toc75 (Wallas et al., 2003). During this process, the M-domain of Toc159 seems to interact with the G-domain of Toc34.

psToc34 functions as a GTP-dependent receptor for preproteins, and is controlled by protein phosphorylation in in vitro experiments (Sveshnikova et al., 2000): GTP-bound Toc34 shows higher affinity for phosphorylated preproteins. This binding stimulates GTP hydrolysis, which causes the release of GDP and the precursor protein (Jelic et al.,
In Arabidopsis, atToc33 may be phosphorylated whereas atToc34 is not (Jelic et al., 2003). However, mutation of the atToc33 protein phosphorylation site in vivo, does not seem to have any effect on protein import (Aronsson et al., 2006).

In Arabidopsis, the knock out mutant of atToc33, the ortholog of psToc34, was generated from a T-DNA insertion within the atTOC33 gene. This mutant, called ppi1, has pale green leaves during the first weeks of development, thereafter, this phenotype disappears (Jarvis et al., 1998). This indicates that atToc33 may be involved during early stage of plant development and that atToc34 can partially complement the loss of Toc33 expression (Jarvis et al., 1998). ppi1 mutant plants also show a defect in expression and import of photosynthetic proteins (Kubis et al., 2003) which suggests that atToc33 is mainly involved in import of photosynthetic proteins.

On the other hand, Arabidopsis plants that do not express atToc34, (this mutant is named ppi3) due to T-DNA insertion, have no specific phenotype apart from delayed root growth. These results confirm the hypothesis that atToc33 can substitute atToc34 and that atToc34 may have a more important role for plastid biogenesis in roots (Constan et al., 2004). The double mutant ppi1/ppi3 is embryo lethal. Therefore, it can be concluded that the function of these two proteins is essential for protein import (Constan et al., 2004).

atToc33 and atToc34 antisense Arabidopsis plants were created. The phenotype for antisense atTOC33 plants is the same as ppi1 plants, but antisense atTOC34 plants have a paler phenotype compared to the wild type when seedlings are 4 days-old, but then the phenotype disappears to get similar to wild type plants (Gutensohn et al., 2000). One possible explanation is that antisense atTOC34 plants may have an effect on atTOC33 expression during early development (Constan et al., 2004), or that atToc33 and atToc34 specificity for photosynthetic or non-photosynthetic proteins is more complex than initially stated (Gutensohn et al., 2000).

5.3 Specificity of import

As mentioned before, in Arabidopsis, atToc159 seems to be involved in a different pathway than atToc132, atToc120 and atToc90, likely, atToc33 and atToc34 also have a different substrate (preprotein) specificity. atToc159 and atToc33 are thought to be dedicated to import photosynthetic proteins, while the others would import non-photosynthetic proteins. In addition, co-immunoprecipitation and in vitro binding experiments showed that atToc159 is binding preferentially to atToc33 rather than atToc34 (Ivanova et al., 2004). On the other hand atToc132 and atToc120 bind equally to atToc33 and atToc34. It is proposed that by competition, atToc159 binds preferentially to atToc33, which leads atToc132 and atToc120 to bind more likely to atToc34.
Fig. 5: An alternative pathway to import non-photosynthetic proteins. (Adapted from Jarvis and Robinson, 1994; Kessler and Schnell, 2006)

In *Arabidopsis*, Toc GTPases are encoded by a small gene family. The different isoforms are involved in at least two different pathways. atToc159 and atToc33 import photosynthetic proteins which constitute the majority of imported proteins. atToc132, atToc120 and atToc34 are involved in the import of non-photosynthetic proteins, also called “housekeeping” proteins. Afterwards, preproteins are imported through the Tic complex. Whether the components of this Tic complex are identical or not for the import of photosynthetic and housekeeping proteins is not known.
These data suggest at least two different pathways of protein import depending on the nature of the substrate (Fig. 5). Photosynthetic proteins could be imported by the Toc complex composed of atToc159, atToc33 and atToc75. Import of non-photosynthetic proteins could rely on atToc132/atToc120 and atToc75. Whether different Tic complexes exist depending on the imported protein still remains to be deciphered (as reviewed by Kessler and Schnell, 2006).

6 Two models for preprotein import

GTP is believed to play an important role in protein translocation due to the presence of the two Toc GTPases, Toc159 and Toc34. Moreover, in vitro import and precursor protein binding experiments require GTP (Olsen and Keegstra, 1992; Kessler et al., 1994). Also, heterodimerisation of Toc159 and Toc34 is favoured in the presence of GDP (Weibel et al., 2003), which confirms the role of GTP hydrolysis in this complex (Kessler et al., 1994; Young et al., 1999).

However, the mechanisms of protein import through the Toc complex and the exact function of the Toc GTPases remain unknown and currently, two models are proposed in pea, the initial plant model for Toc import studies (Fig. 6): the “motor” model and the “targeting” hypothesis. These two models acknowledge the importance of GTP during import, but disagree on the primary Toc receptor for precursor protein.

6.1 The motor model

This model (proposed by Becker et al., 2004b) considers that preproteins are targeted to the chloroplast membrane through the guidance complex (May and Soll, 2000) and that the transit peptide of these precursor proteins is phosphorylated. Becker and associates observed that Toc34 binds to phosphorylated proteins whereas Toc159 does not, thus, Toc34 is proposed to be the initial receptor. More precisely, Toc34 in a GTP bound form (Toc34\textsubscript{GTP}) is binding to the phosphorylated C-terminal part of the transit peptide, then, the N-terminal part of the transit peptide binds to Toc159\textsubscript{GTP}. The interaction with the transit peptide stimulates GTP activity of Toc34 (Jelic et al., 2002) and Toc34 dissociates from the preprotein. Subsequently, the transit peptide is dephosphorylated by an unknown phosphatase since dephosphorylation is required for complete translocation into isolated chloroplast (Waegmann and Soll, 1996). Then dephosphorylation induces GTP hydrolysis of Toc159 and transfers the precursor
Fig. 6: Two models of preprotein import in pea chloroplasts (from Kessler and Schnell, 2006)

A. The “Targeting” model proposes that Toc159 is the primary receptor of the preprotein. Toc159 and Toc34 interact via their G-domain which causes GTP hydrolysis. This GTP hydrolysis changes the conformational state of the Toc complex leading to the import of the preprotein.

B. The “motor” model suggests that Toc34 is the primary receptor in a GTP-bound state of the precursor protein that is in a phosphorylated form. This binding to the precursor induces GTP hydrolysis and the preprotein is transferred to Toc159 in a GTP-bound form. Then after dephosphorylation of the precursor, Toc159 plays a motor role and the precursor is translocated through Toc75.
protein through Toc75. Thereafter the Toc complex regains its GTP bound state, following GDP-GTP exchange, to accept another round of preprotein import.

In this model, Toc159 is a GTP driven motor protein that translocates the preprotein to the intermembrane space (Schleiff et al., 2003b). According to the stoichiometry of the Toc complex of 1/4/4 between Toc159/Toc75/Toc34 (Schleiff et al., 2003b), Toc159 is proposed to be located in the center of this multimeric complex.

However, the fact that phosphorylation is dispensable for protein import (Nakrieko et al., 2004), and that the M-domain alone of Toc159 can partially complement the defect in protein import in the ppi2 mutant (Lee et al., 2003), indicates that this model may not correspond exactly to the import of preproteins in planta. Moreover, some data show that GTP is required early during the import process but not during precursor protein translocation to the intermembrane space (Young et al., 1999; Kessler and Schnell, 2006).

6.2 The targeting model

In this model, Toc159, in its cytosolic form, is thought to be the primary receptor of precursor proteins (Hiltbrunner et al., 2001b). Precursor proteins bind to the G-domain of Toc159GTP but the M-domain of Toc159 may also play a role in this binding process (Smith et al., 2004, Lee et al., 2003).

Toc159 would then interact and bind to Toc34, resulting in insertion of Toc159 to the membrane and GTP hydrolysis. This scheme is supported by the fact that Toc159 and Toc34 are likely to heterodimerise in a GDP-bound conformation (Sun et al., 2002) and that stable association of atToc159 to atToc33 requires the GDP conformation in vitro (Smith et al., 2002). Besides, the activity of both GTPases is required for proper insertion of Toc159 to the membrane (Bauer et al., 2002; Smith et al., 2002; Wallas et al., 2003)

The heterodimer formed by Toc159GDP and Toc34GDP may acquire, due to its GDP-bound form, an adapted conformation to allow the precursor protein translocation through the protein channel, Toc75. Once the precursor is translocated, the Toc GTPases can regain their initial GTP-bound state, and Toc159 is released into the cytosol.

However, this model is not universally accepted because the existence of a soluble form of Toc159 is not widely accepted: some researchers consider that soluble Toc159 is an artefact due to membrane disruption during fractionation experiments (Becker et al., 2004). More in vivo experiments would be necessary to confirm or infirm the targeting model as well as the motor hypothesis.
To resume, GTPase activity is important in both models. In the “motor” model, GTP hydrolysis is involved in precursor protein translocation from one Toc component to another. In the “targeting” hypothesis GTP hydrolysis is involved in the “stability” of the Toc complex, allowing Toc159 insertion to the membrane and interaction of the two GTPases.

7 Aims of the thesis

The exact mechanisms that involve the two Toc GTPases of the Toc complex during import remain unknown. Still, many experiments on in vitro interaction of the two GTPases Toc159 and Toc34 showed that these proteins dimerise via their G-domain (Hiltbrunner, et al., 2001; Sun et al., 2002; Smith et al., 2002; Bauer et al., 2002; Wallas et al., 2003; Weibel et al., 2003). But heterodimerisation of Toc159 and Toc34 still has not been demonstrated in vivo, nor have been observed stable heterodimers (Li et al., 2007).

Today different methods are available to study protein-protein interaction in vivo (as reviewed by Lalonde et al., 2008) many of them are based on the reconstitution of a function of the two halves of a split protein.

The yeast two hybrid system (Fields and Song, 1989) is a powerful heterologous system that is quite efficient to screen many potential candidates but cannot be applied to study membrane proteins. In this case the yeast split ubiquitin assay (Johnsson and Varshavsky, 1994; Stagljar et al., 1998) is more appropriate as it is suitable for studies on membrane and soluble proteins and allows to detect transient interactions. These two methods are mainly used to detect direct protein interactions. To detect secondary interaction within a complex, affinity purification of complexes followed by identification of the proteins in the different complexes by mass spectroscopy (MS) (as reviewed by Miernyk and Thelen, 2008) is a good alternative.

In order to gain more insight into protein import into chloroplasts in the plant itself, one of the goals of my work is to study the interaction between the Toc GTPases using a system for the detection of protein-protein interactions in vivo; for that purpose a system based on the yeast split-ubiquitin in Arabidopsis and Tobacco protoplasts was developed.

It is very likely that still unknown proteins are involved in the mechanism of the Toc complex. To identify such partner proteins, an affinity purification using in vivo tagging followed by MS is a suitable method. For that purpose, the Tandem affinity purification tag (TAP tag) method (Rigaut et al., 1999) was chosen to isolate Toc GTPases and associated complexes from Arabidopsis. Consequently, the second aim of this thesis is
to confirm existing partners of atToc33 and identify new ones with the TAP-tag method.
Results and Discussion

Chapter 1

Use of plant split-ubiquitin to study the interaction between Toc33 and Toc159 GTP-binding domains in vivo

1 Presentation of the system

Ubiquitin is a highly conserved protein of 76 amino acids that is present is all eucaryotes. It is often found covalently attached to other proteins to form monoubiquitin or polyubiquitin chains. Polyubiquitin chains are formed by an isopeptide linkage between the residue at the carboxy-terminus Gly-76 of ubiquitin and lysine residues of proteins ubiquitin is linked to. This chain is a degradation signal for the 26S proteasome (Hershko et al., 2000). The targeted protein is degraded by a proteasome-related protease. During this process, ubiquitin is released from its substrate by deubiquitinating enzymes (DUBs) to be recycled. DUBs are divided into two groups: smaller (10-30kDa) ubiquitin C-terminal hydrolase (UCH) and larger (50-300kDa) ubiquitin specific proteases (UBPs). Besides its function in proteasome-mediated proteolysis, ubiquitin serves as a multi-purpose regulatory modification and is known to have a role during ribosome biogenesis (Finley et al., 1989; Schnell and Hicke, 2003; Stavreva et al., 2006).

In the split ubiquitin system, ubiquitin is spliced and expressed in two separate parts, an N-terminal part (termed Nub, consisting of amino acids 1-37) and a C-terminal part (Cub, amino acids 35-76) fused to a gene coding for a reporter protein (Johnsson and Varshavsky, 1994; Stagljar et al., 1998) (Fig.1). Proteins of interest are fused either to Nub or Cub. If the two proteins interact, the two halves of ubiquitin are brought into close proximity and a quasi-native ubiquitin is reconstituted and recognised by UBPs, resulting in the cleavage of the Cub fusion and the release of the reporter protein.

However, Nub and Cub can reassemble quite easily. To limit unspecific reassociation, two mutations were engineered into Nub in position 13; at this position wild type Nub carries an isoleucine. If isoleucine is replaced by an alanine (NubA) or a glycine (NubG), affinity of the mutated Nub and Cub is lower, (the lowest affinity to Cub is with NubG, Johnsson and Varshavsky, 1994).
Fig. 1 Schematic representation of the split ubiquitin system
Ubiquitin is split in N-terminal (Nub) and C terminal halves (Cub). Each half is fused to a protein of interest. If proteins interact, ubiquitin is reconstituted and recognised by ubiquitin-specific proteases (UBPs). The reporter gene is then cleaved. In the yeast system, the reporter is a nutritional marker or a transcription factor. In the plant system, GFP was chosen as reporter protein.

Fig. 2 Schematic representation of the Toc GTPase proteins atToc159 and atToc33
atToc159 and atToc33 have conserved GTP binding-domains (G-domain, shown in dark grey). In addition, atToc159 has an N-terminal acidic domain (A domain) and a C-terminal membrane anchoring domain (M domain). atToc33 has a short C-terminal hydrophobic transmembrane sequence (TM). In this work, the coding sequence for the G-domain only of atToc159 (Toc159G, Toc159728-1093) was introduced into the different constructs. The atToc33 constructs used contain the coding sequence for G-domain (Toc33G, Toc331-265) or for the entire protein (Toc33).
The yeast split-ubiquitin system has been successfully applied for the study of numerous protein-protein interaction pairs as well as for genome-wide library interaction screens (Lehming, 2002; Miller et al., 2005). Proteins of higher eukaryotes were among the proteins tested, including several, mainly plasma membrane located, plant proteins (Bregante et al., 2008; Deslandes et al., 2003; Ihara-Ohori et al., 2007; Ludewig et al., 2003; Obredlik et al., 2004; Orsel et al., 2006; Pandey and Assmann 2004; Park et al., 2005; Pasch et al., 2005; Reinders et al., 2002a; Reinders et al., 2002b; Schulze et al., 2003; Tsujimoto et al., 2003; Yoo et al., 2005). One disadvantage of the yeast split-ubiquitin system for the study of plant protein interactions is the lack of plant specific factors which might influence the interaction, as observed for the protoplast two-hybrid system (Ehlert et al., 2006). Moreover, for the study of chloroplast outer membrane proteins, yeasts are lacking plastid organelles.

2 Suitability of the yeast split-ubiquitin system to study Toc GTPase interactions

2.1 Toc GTPase interaction studies

Studying Toc GTPases interaction in a yeast system was already successfully attempted in our lab; interaction among the G-domains of atToc33 and atToc159 was demonstrated using the yeast two hybrid system (Dr Birgit Agne, unpublished data). In this thesis it was determined whether Toc159G and Toc33G interacted in the yeast split-ubiquitin system. In the following constructs we engineered, Toc159G corresponds to amino acids 728-1093 of atToc159 and Toc33G to amino acids 1-265 of atToc33 (Fig.2).

Split-ubiquitin fusion constructs were generated by replacing STE14 or PEX11 in the constructs STE14-Cub-RURA3 (Wittke et al., 1999) or Nub-PEX11 (Eckert and Johnsson, 2003) by Toc33G or Toc159G, respectively (Fig.3A). To allow for subsequent Western Blot analyses, two hemagglutinin epitope (HA) tags were introduced upstream of Cub and a Myc epitope downstream of Nub. Constructs encoding for Nub-Pex11p and Pex11p-2HA-Cub-RUra3p were used as a positive control in our experiments as these two fusion proteins were shown to homodimerise using yeast split ubiquitin system (Eckert and Johnsson, 2003). Originally, the arginine-URA3 (RURA3) element was designed to serve as metabolic marker for the interaction between the Nub- and Cub-fusion proteins in growth assays. If the two proteins of interest interact, RUra3p is rapidly degraded by the enzymes of the N-end rule pathway (Varshavsky, 1997), preventing the yeast to grow on media lacking uracil, but allowing them to grow on a 5-Fluoroorotic acid (FOA) supplemented media (Wittke et al., 1999). However, in this thesis, the interactions were monitored by Western blotting.
Fig. 3 Yeast split-ubiquitin interaction of Toc159G and Toc33G.
A. Yeast cells were co-transformed with different combinations of Nub and Cub constructs (a-c). B. Total cellular protein was extracted and analysed by Western blotting using antibodies against the Myc or the HA epitope tag to detect Nub-myc-Toc159G or the Cub-fusion proteins, respectively. Co-expression of Nub-Pex11p and Pex11p-2HA-RUra3p (a) or Nub-myc-Toc159G and Toc33G-2HA-Cub-RUra3p (b) lead to partial cleavage of the RUra3p reporter, indicating interaction of these protein pairs. No cleavage was observed upon co-expression of Nub-Pex11p and Toc33G-2HA-Cub-RUra3p (c). C. Estimation of reporter gene cleavage via chemiluminescence quantification of three independent experiments.
Yeast cells (strain JD53) were co-transformed with the different constructs as shown in Fig.3A. Equal amounts of cellular protein of the transformants were subjected to Western blot analysis with anti-Myc (α-Myc) antibodies to check the presence of Nub-Myc-Toc159G; and with α-HA antibodies to detect the cleavage of the Cub-fusion proteins as indicator of the interaction (Fig.3B). Cleavage (partial) of the Toc33G-Cub fusion protein was observed (Toc33G-Cub-HA, 38kDa) when it was expressed in the presence of Nub-Myc-Toc159G (Fig.3C lane b) whereas no cleavage was observed upon co-expression with a Nub fusion of the peroxisomal protein Nub-Pex11p (Fig.3C, lane c) only a single band corresponding to the Toc33G-Cub-HA-RURA3 (73kDa) fusion protein was detected (negative control). In the positive control experiment, Nub Pex11p fusion protein induced cleavage of Pex11p-2HA-Cub-RUra3p (Fig.3C, lane a), consistent with Pex11p homodimerisation (Eckert and Johnsson, 2003). This cleavage is incomplete like it was for the interaction Nub-myc-Toc159G with Toc33G-Cub-HA-RURA3.

These results suggest that split-ubiquitin is suitable to study Toc protein interactions in yeast and encouraged us to establish the system in plants. In parallel, more experiments on Toc GTPase-precursor protein interactions were carried out in the yeast system as shown below.

2.2 The effects of mutations on Nub upon Toc GTPase

As mentioned at the beginning of the result part, it is possible to mutate one amino acid of ubiquitin (Ile13) to avoid unspecific assembly of the two ubiquitin parts. In order to investigate the effect of Nub mutations upon the interaction of Toc159G and Toc33G, the interaction experiment (Fig.3) was repeated using Nub wt (Fig.4A construct a), NubG (Fig.4A,b), and NubA (Fig.4A,c). While there is cleavage when co-transforming yeast cells by Nub-Myc-Toc159G and Toc33G-2HA-Cub-RURA3 (Fig.4B lane a), this cleavage pattern disappears when using the mutated form of the Nub half (Fig.4B lanes b and c).

2.3 Detection of the interaction between Toc159G and a chloroplast precursor protein

Interaction between Toc159G and a chloroplast precursor protein was also studied in the yeast split ubiquitin system. The coding sequence of the small sub-unit of RuBisCO in its precursor form (pre-ats1B, amino acids 1-181) and mature form (ats1B, amino acids: 55-181) were fused to the Cub-HA-RURA3 cassette (Fig.5A). Yeast cells were co-transformed with Nub-Myc-Toc159G and preats1B-HA-Cub-RURA3 (Fig.5A,b) or
Fig. 4 Yeast split-ubiquitin interaction of Toc159G and Toc33G using wt Nub and NubG (replacement Ile13 by Gly) or NubA (Ile13-Ala).

A. Yeast cells were co-transformed with different combinations of Nub and Cub constructs (a-c). B. Total cellular protein was extracted and analysed by Western blotting using antibodies against the Myc or the HA epitope tag to detect Nub-myc-Toc159G or the Cub-fusion proteins, respectively. Co-expression Nub-myc-Toc159G and Toc33G-2HA-Cub-RUra3p using Nub wt (a), lead to partial cleavage of the RUra3p reporter, indicating interaction of these protein pairs. No cleavage was observed upon co-expression of NubG-myc-Toc159G and Toc33G-2HA-Cub-RUra3p (b) or NubA-myc-Toc159G and Toc33G-2HA-Cub-RUra3p (c). C. Quantification of reporter gene cleavage via chemiluminescence detection of three independent experiments.
Fig. 5. Yeast split-ubiquitin interaction of Toc159G and the small subunit of RuBisCO in its mature (ats1B) and precursor form (preats1B).

A. Yeast cells were co-transformed with different combinations of Nub and Cub constructs (a-f). B. Total cellular protein was extracted and analysed by Western blotting using antibodies against the Myc or the HA epitope tag to detect Nub-myc-Toc159G or the Cub-fusion proteins, respectively. Co-expression of Nub-Toc159G and ats1B-2xHA-RUra3p (d) or Nub-myc-Toc159G and preats1B-2xHA-Cub-RUra3p (b) lead to little cleavage of the RUra3p reporter, indicating interaction of these protein pairs. No cleavage was observed upon co-expression of Nub-Pex11p and preat1B-2xHA-Cub-RUra3p (a) or Nub-Pex11p and at1B-2HA-Cub-RUra3p (c). No cleavage was observed when experiment was repeated using mutated NubG (replacement of Ile13 by Gly) (e and f). C. Quantification of reporter gene cleavage via chemiluminescence detection of three independent experiments.
ats1B-HA-Cub-RURA3 (Fig. 5A,d). The interaction was determined by Western blot. Weak signals corresponding to the cleaved RURA3 were detected (Fig. 5B lanes b and d). As negative controls yeast cells were co-transformed with Nub-Pex11p together with the preats1b- (Fig. 5A constructs a) or ats1B-HA-Cub-RURA3 (Fig. 5A,c) fusion proteins. No cleavage was observed.

The interaction between Nub-Myc-Toc159G and preats1B or ats1B-HA-Cub-RURA3 was repeated using the mutated NubG (Fig. 5A,e,f). As already observed for NubA/-NubG-Myc-Toc159G and Toc33-Cub (Fig. 4B), no cleavage was detected using the mutated Nub-fusions (Fig. 5B lanes e and f).

3 Protein interaction analysis in Arabidopsis protoplasts

For the plant split-ubiquitin system we used plant ubiquitin AtUBQ11 (At4g05050.1) instead of the yeast ScUBI4 and enhanced green fluorescent protein (GFP) as reporter protein. AtUBQ11 is 97% identical to yeast ubiquitin differing from S. cerevisiae Ubi4p by only two amino acids exchanges (S28A, S57A). The N- and C-terminal ubiquitin parts were defined as in yeast, Nub consisting of amino acids 1-37, Cub of amino acids 35-76. Constructs were engineered in the pCL60 vector (Bauer et al. 2000), containing a CaMV 35S promoter and a nopaline synthase (nos) terminator. A HA epitope tag was included in the Cub constructs for subsequent Western blot analysis. Isolated Arabidopsis protoplasts were transformed with constructs encoding for Nub alone, Nub-Toc159G and Toc33G or full length Toc33 fused to HA-Cub-GFP in different combinations (Fig. 6A). The GFP reporter protein of the Cub constructs allowed assessing the protoplast transformation efficiency (around 30% in most of the experiments) by confocal microscopy (Fig. 6B).

Western blots were performed on protein extracts of transformed protoplasts using α-HA antibodies to determine whether cleavage had occurred (Fig. 6C, lower panel) and α-Toc159G antibodies to monitor Nub-Toc159G expression (Fig. 6C, upper panel). When Nub-159G and Toc33G-HA-Cub-GFP were co-expressed, more than 80% cleavage of the GFP reporter was observed (Fig. 6C lane b). In a control experiment in which Nub alone was co-expressed together with Toc33G-HA-Cub-GFP, non-specific cleavage in the range of 40% of the GFP reporter gene was observed (Fig. 4C lane a).

One of our goals, using this method in plant, is to study Toc GTPase interactions in their target membrane. Therefore we performed the same experiment using a Cub construct containing the full length cDNA coding for Toc33 inclusive its C-terminal transmembrane helix (Toc33-HA-Cub-GFP). Around 80% cleavage of the GFP reporter was observed upon co-expression with Nub-159G (Fig. 6C lane c) and around 30 to 40% upon co-expression with Nub alone (Fig. 7 lane c and Fig. 8 lane a).
**Fig. 6** Plant split-ubiquitin interaction for Toc159G and Toc33

A. *Arabidopsis* protoplasts were co-transformed with Nub and Cub constructs as indicated (a-c).

B. Use of the GFP reporter to assess protoplast transformation efficiency via confocal microscopy. Due to the background cleavage all Cub-GFP-fusions gave the same fluorescence pattern as exemplified here for Toc33-HA-Cub-GFP (bar = 5µm). C. Total proteins were extracted and analysed by Western blotting using antibodies raised against Toc159G and anti-HA to check for the presence of Nub-Toc159G and the HA-tagged Cub-fusion proteins, respectively. Co-expression of Nub-Toc159G with Toc33G-Cub-GFP (b) or full-length Toc33-Cub-GFP (c) results in almost complete cleavage of the GFP reporter indicating interaction. The control experiment (a) reveals background cleavage of the reporter gene. D. Experiments (a) and (b) were carried out in tobacco protoplasts and gave similar cleavage patterns. The graphs below show the results of chemiluminescence quantification of three independent experiments.
3.1 The plant split-ubiquitin system is applicable to other plant species

To test whether the plant split-ubiquitin system could be used to study protein interactions in other plant species, the Nub-Toc159G and Toc33G-HA-Cub-GFP constructs were transformed into isolated tobacco protoplasts (Fig. 6D). In a control experiment in which Nub alone was co-expressed together with Toc33G-HA-Cub-GFP in tobacco protoplasts, non-specific cleavage in the range of 50% of the GFP reporter occurred. When Nub-Toc159G and Toc33G-HA-Cub-GFP were co-transformed, cleavage of the reporter gene rose to around 90%.

Taken together, the increase in cleavage of the Toc33-Cub fusion proteins in the presence of Nub-Toc159G indicated interaction of these proteins \textit{in planta}.

3.2 The mutated Nub$^G$ is not suitable to study Toc protein-protein interactions

In order to limit the background cleavage observed with non interacting proteins, the Nub$^G$ mutation was introduced in the plant split-ubiquitin experiments (Fig.7). The background cleavage is still present whether protoplasts are co-transformed with Nub and Toc33-HA-Cub-GFP (Fig.7 lane b) or with Nub$^G$ and Toc33-HA-Cub-GFP (Fig.7 lane c). However the complete cleavage pattern observed when protoplasts transiently express Nub-Toc159G and Toc33-HA-Cub-GFP (Fig.7 lane e) is strongly reduced when co-expressing the mutated Nub$^G$ Toc159G and Toc33-HA-Cub-GFP (Fig.7 lane d). Using Nub$^G$ gave exactly the same result as the one observed with the yeast system, interaction using Nub$^A$ mutation was not tested. This result indicates that only wt Nub is suitable to study Toc protein interactions. One possible explanation could be that Toc proteins interact only transiently and a mutated ubiquitin is not suitable to reveal the interaction, because the mutation in the Nub may slow down reconstitution.

As a result of background cleavage, the GFP reporter served only as an indicator of the protoplast transformation rate, but could not be exploited to localize the Cub fusion proteins in the absence of an interacting protein. The transformation of all Cub-GFP fusion proteins used in this study resulted in essentially the same cytosolic fluorescence pattern as can be seen in Fig. 6B. The high rate of non-specific cleavage in the plant split-ubiquitin system is in contrast to the very low rate of background cleavage observed in the yeast split-ubiquitin assays (Fig.3) and could be due to a comparatively high activity of \textit{Arabidopsis} UBPs (Yan et al., 2000).
Fig. 7 Plant split-ubiquitin interaction among Toc159G and Toc33 using wtNub and NubG.

A. Arabidopsis protoplasts were co-transformed with Nub and Cub constructs as indicated (a-e). B. Total proteins were extracted and analysed by Western blotting using antibodies raised against Toc159G and anti-HA to check for the presence of Nub-Toc159G and the HA-tagged Cub-fusion proteins, respectively. Co-expression of NubToc159G with Toc33-Cub-GFP (e) results in almost complete cleavage of the GFP reporter indicating interaction. This complete cleavage disappears with coexpression of NubG Toc159G with Toc33-Cub-GFP. The control experiments (b and c) reveal partial background cleavage of the reporter gene, as well as expression of Toc33-Cub-GFP alone (a). C. The graphs below show the results of chemiluminescence quantification of three independent experiments.
3.3 General suitability of the plant split-ubiquitin system to analyze protein interactions

To test whether the plant split-ubiquitin system is useful for protein pairs other than Toc33 and Toc159 constructs encoding Nub and Cub fusions to an *Arabidopsis* homolog of yeast Pex11 were engineered. Five PEX11 homologs were identified in *Arabidopsis* (atPEX11a-e), all representing peroxisomal membrane proteins involved in peroxisome proliferation (Lingard and Trelease, 2006; Orth *et al*., 2007). Two out of these five homologs, atPEX11c and atPEX11e have been demonstrated to partially complement the *S. cerevisiae* pex11 null mutant (Erdmann and Blobel, 1995) indicating a conserved function in peroxisome biogenesis and similar interaction patterns (Orth *et al*., 2007). atPex11e was selected as the model protein and constructs were engineered encoding Nub-Pex11e and Pex11e-HA-Cub-GFP (Fig.8A) in order to test for atPex11e self-interaction in our newly developed plant split-ubiquitin system. Pex11e was not expected to interact with either Toc33 or Toc159. Therefore Pex11e may also be useful as a negative control for Toc GTPase interactions. Co-expression of Nub-Pex11e and Pex11e-HA-Cub-GFP in isolated tobacco protoplasts gave around 85% reporter GFP cleavage (Fig.8B lane d) indicating interaction whereas co-expression of Pex11e-HA-Cub-GFP with either Nub alone (Fig.8B lane c) or Nub-Toc159G (Fig.8B lane e) resulted in only 30% and 40% of cleavage. Similarly, co-expression of Nub-Pex11e with Toc33-HA-Cub-GFP resulted in around 45% cleavage of the GFP-reporter (Fig.8B lane f). Thus, the cleavage observed when co-expressing a Toc GTPase with Pex11e is similar to the one observed upon co-expression of Nub and Toc33-HA-Cub-GFP (Fig.8B lane a) and therefore most probably unspecific.

3.4 Toc protein-protein interactions in the protoplast cytosol

To test whether the fusions of the membrane proteins Toc33 and Pex11e insert into membranes the experiments depicted in Fig.8 were repeated with an additional cell fractionation step. Extracts of transformed tobacco protoplasts were centrifuged at 100000xg to separate soluble proteins (Fig.9, S “soluble”) from membrane proteins (Fig.9, P “pellet”). Western blot analysis with α-HA revealed that both the uncleaved and cleaved form of full length Toc33HACub-GFP were predominantly located in the soluble fraction (Fig.9 lanes a,b, S). Only upon co-expression of Nub-Toc159G, a small portion of cleaved Toc33-HA-Cub was detected in the 100’000xg pellet fraction (Fig.9 lane b, P). These data suggest that the C-terminal HA-Cub-GFP fusion prevents insertion of Toc33 into the membrane, and that only upon cleavage of the GFP, Toc33 membrane insertion is possible. Therefore, the interaction we observed between Nub-Toc159G and full length Toc33-HA-Cub-GFP in the plant split-ubiquitin system most likely occurs in the protoplast cytosol. The uncleaved and cleaved fusion of the second
Fig. 8 Toc and Pex protein interactions in the plant split-ubiquitin system

A. Tobacco protoplasts were co-transformed with Nub and Cub constructs. B. Total proteins were extracted and analysed by Western blotting using antibodies raised against Toc159G to check for the presence of Nub-Toc159G or anti-HA for the Cub-fusion proteins. C. Interacting protein pairs result in almost complete cleavage of GFP [NubToc159G and Toc33-HA-Cub-GFP (b), NubPex11e and Pex11e-HA-Cub-GFP (d)], non-interacting protein pairs result in partial background cleavage of the reporter gene [Nub and Toc33-HA-Cub-GFP (a), Nub and Pex11e-HA-Cub-GFP (c), NubToc159G and Pex11e-HA-Cub-GFP (e), NubPex11e and Toc33-HA-Cub-GFP (f)].
Fig. 9 Membrane association of full length Toc33 and Pex11e in plant split-ubiquitin assays.

A. Tobacco protoplasts were co-transformed with Nub and Cub constructs. B. Protoplasts were lysed and separated into soluble and pellet fraction by centrifugation at 100,000 x g for 1h. Equal amounts of protein of non-treated protoplasts (N), soluble (S) and pellet (P) fraction were analysed by immunoblotting using antibodies against Toc159G, the HA epitope, Toc75 and phosphoribulokinase (PRK). Toc75 and phosphoribulokinase served as membrane and soluble marker, respectively. Co-expression of Nub-Toc159G with full length Toc33-HA-Cub-GFP (b) or Nub-Pex11e with Pex11e-HA-Cub-GFP (d) results increased cleavage. Uncleaved and cleaved form of the Toc33-HA-Cub fusion (a, b) are mainly present in the soluble fraction suggesting inhibition of Toc33 membrane insertion by the C-terminal fusion part. In marked contrast, uncleaved and cleaved fusions of the integral membrane protein Pex11e (c, d) are both located in the pellet fraction.
membrane protein tested, Pex11e, were mainly located in the insoluble fraction (Fig. 9 lanes c,d). This indicates that in contrast to Toc33, membrane insertion of Pex11e is not affected by the C-terminal fusion part. Thus, in the plant split-ubiquitin system it is feasible to study interactions of membrane associated Pex11e.

4 Discussion

4.1 Comparison of the plant split-ubiquitin system with other available in vivo protein-protein interaction assays

In response to the emerging need for in vivo protein-protein interaction data several in vivo assay systems have been developed in the recent past. Most of these are based on protein fragment complementation and have been demonstrated to be applicable to plant cells as well. Examples are bimolecular fluorescence complementation (BiFC) (Bhat et al., 2006; Hu et al., 2002; Kerppola, 2008), fluorescence resonance energy transfer (FRET) (Miyawaki et al., 1997), split-mDHFR (murine dihydrofolate reductase) (Subramaniam et al., 2001) or split-luciferase (Fujikawa and Kato, 2007) protein complementation assays and the protoplast two-hybrid (P2H) system (Ehlert et al., 2006).

The availability of different protein-protein interaction assay systems is advantageous as each has his inherent limitations and might not be applicable to all pairs of interacting proteins. Membrane receptor GTPases at the chloroplast outer surface are presumed to undergo short-lived and dynamic interactions with chloroplast precursor proteins and among themselves. For their study in vivo, it is necessary to use a system that allows the analysis of transient protein-protein interaction at the cytosolic face of organelles. For that purpose we adapted the yeast split-ubiquitin system possessing the desired characteristics for plant cells and demonstrate the interaction between Toc159 and Toc33 as well as Pex11self-interaction.

4.2 Toc GTPase heterodimerisation in vivo

In many in vitro studies homo- or heterodimerisation of the G-domains of Toc33 and Toc159 has been observed (Bauer et al., 2002; Hiltbrunner et al., 2001b; Reddick, et al., 2007; Smith et al., 2002; Sun et al., 2002; Weibel et al., 2003; Yeh et al., 2007). Working with recombinant or in vitro translated proteins, stable homodimers of Toc159 and Toc33 are much more easily obtained than heterodimers leading to the assumption
that Toc159 and Toc33 do not form stable heterodimers or that heterodimers are formed only transiently in vivo (Li et al., 2007).

A short-lived interaction among Toc159 and Toc33 fits well with a model of a highly dynamic, nucleotide dependent Toc GTPase cycle in chloroplast protein import. In our lab we have demonstrated for the first time the heterodimerisation among the Toc GTPases Toc159 and Toc33 in three different in vivo interaction assay systems: the yeast two-hybrid system (previous work done by Dr Birgit Agne), the yeast split-ubiquitin system and the plant split-ubiquitin system, the latter especially developed for this purpose. Surprisingly and in contrast to in vitro studies mentioned above, Dr Birgit Agne could not observe Toc33G-Toc33G or Toc159G-Toc159G homodimerisation in the yeast two-hybrid system. This is why we haven’t studied homodimerisation by split-ubiquitin yet. However, homodimerisation in plant split ubiquitin system, will have to be tested in the future, moreover, our results supply evidence that heterodimerisation indeed occurs in vivo.

The yeast two-hybrid interaction is an excellent tool to screen for mutations altering the binding properties of Toc159G for Toc33G and vice versa. The resulting mutations could be subsequently assayed in planta by split-ubiquitin. By cell-fractionation via ultracentrifugation we demonstrated that, the interaction observed in the plant split-ubiquitin system among Toc159G and full-length Toc33, occurred in the cytosol and not at the chloroplast membrane. Apparently, the C-terminal GFP fusion interfered with Toc33 membrane insertion. The data suggest that the C-terminus of Toc33 must be freely accessible for membrane insertion. This is supported by the insertion of a small portion of Toc33-HA-Cub after cleavage. By switching orientation of the fusion parts and including full length Toc159 the system is currently improved for the study of Toc GTPase interactions at the chloroplast surface.

4.3 Interaction with precursor proteins

In the yeast split-ubiquitin system, we could demonstrate the interaction between Toc159G and the precursor protein of the small sub-unit of RuBisCO, preats1b. Surprisingly, Toc159G seems to interact also with the mature protein. But the cleavage observed is quite low, less than 20% (Fig.5), compared to the 70% of cleavage observed with the presence of Nub-Toc159G and Toc33G-Cub fusion proteins (Fig3). This very weak cleavage could be due to very short-lived interaction or a problem of mistargeting into mitochondria. Indeed in organisms that do not have chloroplasts such as yeast cells, chloroplastic proteins like RuBisCO (Rudhe et al., 2002) and ferredoxin (Hirohashi et al., 2001) can be mistargeted into mitochondria (Hurt et al., 1986). To test this hypothesis, we could test in the near future the interaction between Toc proteins and preats1b in the plant split-ubiquitin system.
4.4 atPex11e self-interaction

By means of immunoprecipitation the mammalian PEX11 isoforms PEX11α and PEX11β have been shown to homo- but not heterodimerise (Li and Gould, 2003). No such data is available for the Arabidopsis PEX11 family comprising five members (named by letters a to e). The plant split-ubiquitin experiments carried out in this study show in vivo homodimerisation of atPex11e. In the case of atPex11e (in contrast to Toc33-HA-Cub-GFP) the C-terminal Cub-GFP fusion was almost entirely present in the membrane pellet after centrifugation at 100’000xg (Fig.9). The C-terminal GFP did therefore not appear to interfere with membrane insertion. This result demonstrates that plant split-ubiquitin may indeed be applied to determine and analyze interactions between integral membrane proteins. Specifically the plant split-ubiquitin system may be used to test homo- or heterodimerisation among the Arabidopsis PEX11 homologs in vivo.

4.5 Future modification and improvement of the plant split-ubiquitin system

Some improvement of the plant split-ubiquitin system will need to be done, particularly concerning the reduction of background cleavage and the reporter system.

A higher level of background cleavage of the Toc33-Cub-fusion was observed in the absence of an interacting protein in the plant than in the yeast split-ubiquitin assays. This is not due to a higher rate of spontaneous in vivo association of the Nub and Cub fragments in plants as the same level of background cleavage was observed when the Cub-fusion proteins were expressed in the absence of free Nub or Nub-fusion proteins (Fig.7). Possible explanations are that substrate recognition by plant UBPs is less dependent on a complete ubiquitin moiety or that the overall activity of UBPs in plants is higher than in yeast.

The latter appears likely as about twice as many DUBs have been identified in Arabidopsis thaliana compared to S. cerevisiae (Yang et al., 2007). For a further development of the plant split-ubiquitin system reduction of the background-cleavage would be desirable. Potentially, this could be achieved by performing the assays in protoplasts derived from mutant plants in which selected, nonessential UBPs are knocked-out. Furthermore, optimization of assay duration or other parameters could improve the ratio between unspecific background and interaction induced cleavage.

The yeast split-ubiquitin system is available with different reporter molecules. For the study of interactions between membrane proteins a system with a transcription factor reporter was created (Stagljar et al., 1998). In this system the Cub-reporter is fused to a membrane-associated protein and transcriptional activation of a secondary reporter only
occurs upon cleavage and liberation of the primary reporter (transcription factor). A similar system is conceivable for the plant split-ubiquitin system combining it for example with the transcription activation system described by Moore (Moore et al., 1998): using the chimeric transcription factor LhG4 as cleavable reporter fused to a plant membrane protein and protoplasts of transgenic plants with an pOp-β-glucuronidase (GUS) insertion would allow for quantification of interaction via GUS activity assays. This system would be similar to the protoplast two-hybrid system with the advantage that membrane protein interactions could be monitored. A problem in obtaining quantitative data from transient expression experiments in protoplasts is the variability of transformation rates. As in the protoplast two-hybrid system, variant protoplast transformation rates could be corrected by co-expression of the GUS compatible reporter NAN (synthetic neuraminidase gene) (Ehlert et al. 2006).

Thus, this work presented in this manuscript provides an ideal starting point for the optimization and further development of split-ubiquitin system to exploit its full potential in plant systems.

(The results presented on this first chapter have been accepted for publication: Rahim et al., 2008(?))
Chapter 2

Isolation of atToc33 and associated complexes from Arabidopsis.

1 Presentation of the system

The TAP (tandem affinity purification) tag (Rigaut et al., 1999) is composed of two IgG binding domains of Staphylococcus aureus Protein A (ProtA), and a calmodulin binding peptide (CBP) separated by a TEV (Tobacco Etch Virus) protease cleavage site (Fig.1B). These two tags are designed to isolate and purify fusion proteins under native conditions. In this thesis, a modified TAP tag (TAPi) lacking a plant nuclear localization signal initially present in the CBP domain (Rohila et al., 2004) has been employed.

The method consists of two sequential purification steps (Fig.1B). First, the ProtA of the fusion protein is bound to IgG coated-beads. The fusion protein is eluted under native conditions using TEV protease. Second, the eluted protein is incubated with a calmodulin matrix. Elution is performed under native conditions using EGTA.

Alternatively to TEV protease cleavage, protein complexes can be eluted from IgG beads with glycine at pH3. This acidic elution elutes the entire TAP-tag-fusion protein and complexes from the IgG beads.

2 Generation of transgenic plants

The cDNAs of atToc33 and the G domain of atToc33 (Toc33G, aa 1-265) were fused to the coding sequence of the N-terminal TAPi tag (NTAPi) (Fig.1A) and ligated into a binary vector (pCAMBIA) containing the cauliflower mosaic virus 35S promoter, the nopaline synthase terminator and a phosphinotricin (BASTA)-resistance gene for selection of the transformants. Wild type (Col0) and atToc33 knockout mutant ppi l plants were transformed with these two constructs using the floral dip method (Clough and Bent, 1998) (generation T0). This resulted in the following lines: Col0 NTAPiToc33, Col0 NTAPiToc33G, ppi l NTAPiToc33 and ppi l NTAPiToc33G.
Fig. 1 Structure of the TAP tag expression cassette and overview of the different TAP purification steps (from Puig et al., 2001)

A. Schematic representation of the two NTAPI-Toc33/Toc33G constructs.

B. Overview of the TAP protocol. During the first step, TAP-tagged proteins are sequestered by IgG sepharose and released by TEV-protease cleavage. TEV protease–cleaved proteins are then bound to a Calmodulin matrix and eluted using EGTA.
Fig. 2 Analysis of transgenic plants expressing NTAPI-To33 and NTAPI-To33G.
A. BASTA resistant plants were selected and DNA was extracted from 4 weeks-old-plant leaves. 
B. PCR was performed with primers specific for the T-DNA insertion in TOC33 (“ppi1”), as well as for the transgenes NTAPI-To33 and NTAPI-To33G (“NTAPI-To33/To33G”). C. Tap-tagged protein expression was monitored by Western blotting using specific antibodies raised against To33. Each lane corresponds to a different plant line. 25µg of protein were applied per lane.
Wild type Col0 and *ppi1* plants were used as controls.

Phosphinotricin resistant seedlings (generation T1) (Fig.2A) were screened by PCR to detect both the *ppi1* T-DNA insertion (Jarvis *et al*., 1998) and the transgene (Fig.2B). A Western blot using α-atToc33 antibodies was also performed on total proteins extract to check the expression levels of the TAP-tagged Toc33 proteins (Fig.2C). This immunoblot produced a signal at the expected sizes: 54kDa and 49kDa for NTAPI-Toc33 and NTAPI-Toc33G respectively. The strong signal observed in some lanes (*ppi1* NTAPI-Toc33G#2, Col0 NTAPI-Toc33G#1, Col0 NTAPI-Toc33#1) suggests a very high expression level of the fusion protein. T1 generation plants expressing NTAPI-Toc33 or NTAPI-Toc33G according to these Western blots were selected and grown to produce seeds. Resulting T2 phosphinotricin-resistant lines were grown on soil under short day conditions to perform a small-scale purification of the TAP tag complex.

3 Characteristics of plants expressing NTAPI-Toc33/Toc33G fusion proteins

3.1 Phenotype and chlorophyll measurements

Plants grown on soil that expressed NTAPI-Toc33 or NTAPI-Toc33G showed different phenotypes (Fig.3B). *ppi1* NTAPI-Toc33G plants were much paler than Col0 NTAPI-Toc33G plants. *ppi1* NTAPI-Toc33 plants were slightly paler than Col0 NTAPI-Toc33 plants but had a stronger green phenotype compared to untransformed *ppi1* plants. This phenotype was also observed on plants that were grown in vitro (Fig.3A).

Chlorophyll was measured on the four different transgenic lines grown on plates (Fig.3C) using Col0 and *ppi1* plants as controls (*ppi1* mutant plants contain reduced levels of chlorophyll at early developmental stages (Jarvis *et al*., 1998)). These measurements revealed similar chlorophyll amounts in Col0, Col0 NTAPI-Toc33 and Col0 NTAPI-Toc33G plants. The chlorophyll amount in *ppi1* NTAPI-Toc33G plants is similar to that found in *ppi1* plants. In *ppi1* NTAPI-Toc33 plants, chlorophyll is slightly weak compared to Col0 NTAPI-Toc33 plants, but definitely higher than *ppi1* plants. These results suggested that NTAPI-Toc33 complemented the *ppi1* phenotype and defect on chlorophyll levels, whereas NTAPI-Toc33G did not.
Fig. 3: Phenotype of plants expressing TAP tagged Toc33
Pictures of 3 weeks old plants growing on plate (A) or on soil (B) under short day conditions.
Chlorophyll measurement of plants expressing TAP tagged Toc33 grown in vitro (C).
Fig.4: Immunolocalisation of TAP-tagged Toc33 protein using antibodies raised against atToc33. Protoplasts were fixed and probed with anti-toc33 antibodies. Secondary FITC-coupled antibodies were used for visualisation. The detection of the chromophore fluorescence (FITC) by confocal microscopy, chlorophyll auto-fluorescence, bright-field microscopic images (visible) and merge (from bright-field and fluorescent images) are shown.
3.2 Localisation of the fusion proteins

The *ppi1* phenotype is complemented by the NTAPi-Toc33 fusion protein but not by NTAPi-Toc33G. This could be explained by the fact that NTAPi-Toc33 unlike NTAPi-Toc33G is inserted to the chloroplast outer membrane. To test whether NTAPi-Toc33 is localized in the chloroplast outer membrane, fluorescence immunolocalisation of the fusion proteins NTAPi-Toc33 and NTAPi-Toc33G was carried out. Protoplasts were isolated, fixed and probed with α-atToc33 (Fig.4) and α-IgG antibodies (data not shown). Fluorescence (FITC) was detected by confocal microscopy. Probing with anti-atToc33 gave an identical cytosolic fluorescence pattern for both NTAPi-Toc33 and NTAPi-Toc33G with no apparent difference between the two.

In parallel, high speed (100’000xg) centrifugation of total plant protein extracts was done to separate soluble (SN) and membrane fractions (P) (Fig.5). Both fractions were analysed by SDS-PAGE, transferred to a nitrocellulose membrane and blotted with antibodies raised against components of the Toc-complex: α-atToc75, α-atToc159G, α-atToc159A and α-atToc33.

Toc75, the Toc protein channel, was indeed located in the membrane fraction in all samples tested. Toc159 was mainly present in the insoluble fraction. Some Toc159 was also detected in the soluble fraction and seemed to be more abundant in untransformed *ppi1* plants and of plants expressing Col0 NTAPi-Toc33G. Surprisingly, in Col0 plants expressing NTAPi-Toc33 or NTAPi-Toc33G (Fig.5A left), both NTAPi-Toc33 and NTAPi-Toc33G were detected in the soluble and insoluble fractions. In *ppi1* plants expressing NTAPi-Toc33 (Fig.5 right), NTAPi-Toc33 was detected in the membrane fraction only. In *ppi1* NTAPi-Toc33G plants, NTAPi-Toc33G was present in both soluble and pellet fractions, but mainly in the soluble one.
Fig.5: Localisation of Tap-Tagged Toc33 proteins and complexes by ultracentrifugation

Plant extracts were centrifuged at 100’000g for 1h. Supernatant (SN) and pellet (P) were collected separately and analysed by SDS-PAGE followed by Western blotsting with antibodies raised against atToc75, atToc132 and atToc159A or atToc159G. Protein sizes are indicated in kDa (L: protein molecular weight marker).
A small scale TAP purification was carried out with 1mg of total protein (3-6 leaves). A medium scale TAP purification was carried out with 50 mg of total proteins (10g of leaves). The protein extract was cleared by centrifugation and incubated with IgG coated beads. The protein complexes bound to the beads were washed and eluted with 0.1M glycine pH3 or with TEV protease. Unless specified, all steps were done at 4 ºC.

**Fig.6: Overview of the TAP procedure.**

A small scale TAP purification was carried out with 1mg of total protein (3-6 leaves). A medium scale TAP purification was carried out with 50 mg of total proteins (10g of leaves). The protein extract was cleared by centrifugation and incubated with IgG coated beads. The protein complexes bound to the beads were washed and eluted with 0.1M glycine pH3 or with TEV protease. Unless specified, all steps were done at 4 ºC.
4 Purification of NTAPi-Toc33 and associated proteins

4.1 Small scale purification assays

A small scale purification experiment from a few leaves of Col0 and ppi1 plants expressing either NTAPi-Toc33G or NTAPi-Toc33 was performed as schematized in Fig.6. Proteins were extracted, according to Rensink et al., 1998, in a buffer that contained 0.5% v/v Triton X-100 to extract membrane proteins. Proteins were incubated with IgG-coated beads, washed with buffer, and eluted with glycine pH3. In order to determine the presence of TAP-tagged proteins at each purification step, protein extracts, flow-through, eluates and beads were subjected to Western blot analysis using α-atToc33 antibodies (Fig.7). Both NTAPi-Toc33G and NTAPi-Toc33 proteins showed binding to IgG beads. Protein recovery from IgG beads by acidic elution appeared to be efficient for both TAP fusion proteins.

The same experiment was also subjected to immunoblot analysis using antibodies raised against atToc75, atToc132 and atToc159G to detect whether these Toc proteins co-purified with the TAP fusion proteins (Fig.7). The NTAPi-Toc33 protein co-purified with atToc75, atToc132 and atToc159G both from Col0 and ppi1 backgrounds. atToc159 was also detected in Col0 NTAPi-Toc33G eluates (Fig.7). No other Toc protein tested co-purified with Col0/ppi1 NTAPi-Toc33G eluates.

Col0 NTAPi-Toc33G immunoblots always gave strong signals in the eluates fraction using α-atToc75 and α-atToc132. These were considered unspecific because these signals were present in all blotting experiments, no matter what antibodies were used. This is probably due to the high expression level of NTAPi-Toc33G and the high reactivity of ProtA tag with the antibodies.

4.2 Elution of NTAPi-Toc33 and NTAPi-Toc33G by TEV protease

The same small scale experiment was repeated. This time proteins were incubated with IgG-coated beads, washed and then treated with TEV protease (Fig.6). Aliquots of each purification step were subjected to Western blot analysis.
Fig. 7 Western blot analysis of the Toc proteins associated with NTAPI-Toc33
Total protein (900 µg) (1) of 6 weeks-old plants was incubated with IgG coated beads during 2 ½ hours, the unbound proteins (2) were discarded. The beads were then washed 3 times (3) before elution. The protein bound to the beads were eluted with glycine pH 3 (4), the beads and remaining proteins after elution were extracted in SDS-PAGE sample buffer (5). Samples were analysed by SDS-PAGE followed by Western blotting with antibodies raised against atToc33, atToc75, atToc132 and atToc159G. Protein sizes are indicated in kDa. (L: empty lane).

Fig. 8: NTAPI-Toc33/Toc33G and associated proteins purified from Col0 or ppi1 plants.
Proteins were extracted and purified as described before. Western blots were performed with antibodies raised against Toc33, Toc159, Toc132 and Toc75 (Fig.7). The results are resumed in this figure.
Fig. 9: Purification of Tap-Tagged Toc33 protein complexes and elution by TEV cleavage.
Non transformed plants (A) and plants expressing the TAP-tag fusion protein (B) were used for this experiment. 3g of plants were ground resulting in 6mg of total protein that were bound to the beads (1). Different steps of the purification were loaded on a 12% SDS gel (of each fraction 1/5 of total volume were applied per lane): unbound proteins (2), TEV eluates (3), Boiled beads after TEV elution (4). Presence of Tap-tagged Toc33 was monitored using antibodies raised against Toc33. The purification experiment was performed with Triton or without any detergent (T-). Presence of the different Toc proteins in the eluate were checked with antibodies raised against atToc75, atToc132 and atToc159G. Protein sizes are indicated in kDa. (L: protein molecular weight marker. Upper arrows correspond to NTAPi-Toc33 or NTAPi-Toc33G, Lower arrow correspond to CBP-Toc33 or CBP-Toc33G).
Immunoblot using α-atToc33 antibodies showed shift in molecular mass of 15kDa in the TEV eluates of NTAPi-Toc33G and NTAPi-Toc33 fusion proteins when compared to the acidic elution. This demonstrated that elution from IgG beads using TEV protease is working (Fig.9B). atToc75, atToc132 and atToc159 were also present in NTAPi-Toc33 eluates although the signals corresponding of these three Toc proteins were much weaker in ppi1 NTAPi-Toc33 samples (with a very faint signal for Toc75, Fig.9B arrow). In this experiment, atToc75, atToc132 and atToc159 did not co-purify with Col0/ppi1 NTAPi-Toc33G proteins.

In these “small scale” assays, proteins were extracted in a buffer that contained Triton X-100 detergent to solubilise membrane proteins. NTAPi-Toc33G was expected to be cytosolic, but we previously showed, by ultracentrifugation, that NTAPi-Toc33G was present in both soluble and insoluble fractions (Fig.5). Col0 and ppi1 plants expressing NTAPi-Toc33G were subjected to protein extraction using a Triton-free (T-) extraction buffer. Proteins were isolated as described above. Immunoblot analysis using α-atToc33 antibodies showed that the TAP-fusion protein was bound to and recovered from the IgG beads in the absence of detergent (Fig.9B NTAPi-Toc33G Col0/ppi1 T-).

As a negative control, total proteins from untransformed plants, Col0 and ppi1, were extracted and subjected to TAP tag purification and Western blot analysis (Fig.9A). As expected neither Toc75, nor Toc132 or Toc159 were present in the eluates.

5 Mass spectrometric protein identification

To attempt identification of NTAPi-Toc33 associated proteins, a larger scale TAP purification was performed (Fig.6). Proteins were solubilised from 10g of leaves using the Triton X-100 buffer as described above. IgG affinity chromatography was carried out and the TEV protease eluates were separated by SDS-PAGE and visualized using SYPRO Ruby staining (Fig.10A). Two bands corresponding to the IgG heavy and light chains, at around 50 and 25kda respectively, appeared in the eluates. In addition to these two bands, several other bands were observed (Fig.10A). These bands were excised (Fig.10B) and analysed by mass spectrometry after in gel digestion of the proteins with trypsin. In Col0 and ppi1 plants bands of the same mass as the ones chosen in TAP tag eluates were selected as controls (Fig.10B).

Data received from peptide analysis from each band are presented at the end of this manuscript (see Appendix). The peptide profile showed not many differences whether bands were cut from Col0/ppi1 NTAPi-Toc33 or NTAPi-Toc33G plants (Appendix). A selection of these results is presented in Table1.
Fig. 10 Visualisation of SYPRO Ruby colored gel before trypsin digest

The TEV eluates were analysed by SDS-PAGE followed by SYPRO Ruby staining. Staining is visualized by UV light (A). Control eluates from Col0 and ppi1 were used as negative controls. Bands exclusively present in plants expressing TAP tag constructs were selected to be trypsin digested (B). The white rectangles indicate areas of the gel excised for mass spectrometric analysis.
<table>
<thead>
<tr>
<th>Protein description</th>
<th>Predicted localisation</th>
<th>Accession number</th>
<th>Number of peptides</th>
</tr>
</thead>
<tbody>
<tr>
<td>atToc33</td>
<td>Outer envelope membrane</td>
<td>AT1G02280</td>
<td>114</td>
</tr>
<tr>
<td>atTic110</td>
<td>Inner envelope membrane</td>
<td>AT1G06950</td>
<td>53</td>
</tr>
<tr>
<td>Actin</td>
<td>Cytoskeleton</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cpn60A (CHAPERONIN60 ALPHA)</td>
<td>Chloroplast</td>
<td>AT2G28000</td>
<td>40</td>
</tr>
<tr>
<td>Cpn60B (CHAPERONIN 60 BETA)</td>
<td>Chloroplast</td>
<td>AT1G55490</td>
<td>29</td>
</tr>
<tr>
<td>Hsp93-V</td>
<td>Chloroplast stroma</td>
<td>AT5G50920</td>
<td>50</td>
</tr>
<tr>
<td>Ptac16 (PLASTID TRANSCRIPTIONALLY ACTIVE18); similar to catalytic/ coenzyme binding [Arabidopsis thaliana] similar to Tic62 protein [Pisum sativum]</td>
<td>Chloroplast thylakoid membrane</td>
<td>AT3G18890</td>
<td>4</td>
</tr>
<tr>
<td>Tubulin</td>
<td>Various compartments, cytoskeleton</td>
<td>AT1G04820</td>
<td>7</td>
</tr>
<tr>
<td>EF1α Elongation factor</td>
<td>Cytoskeleton</td>
<td>AT1G07920</td>
<td>43</td>
</tr>
<tr>
<td>Patl1 (PATELLIN 1) Protein targeting and vesicular trafficking</td>
<td>Cytosolic</td>
<td>AT1G72150</td>
<td>13</td>
</tr>
<tr>
<td>DnaJ heat shock family protein</td>
<td>Endomembrane</td>
<td>AT3G62600</td>
<td>3</td>
</tr>
<tr>
<td>cpHSC70-2</td>
<td>Chloroplast stroma</td>
<td>AT5G49910</td>
<td>46</td>
</tr>
<tr>
<td>(Calmodulin 4); calcium ion binding calcium-binding protein, putative</td>
<td>?</td>
<td>AT1G66410</td>
<td>5</td>
</tr>
</tbody>
</table>

|                | ? | AT1G18210 | 6 |

Table 1: Mass spectrometry analysis of TEV eluted proteins from NTAPI-Toc33 purification experiments.
Eluates were loaded on a commercial Biorad 4-12% Bis-Tris gel coloured with SYPRO Ruby. Bands were excised and digested in gel with trypsin. Peptides eluted from the gel piece were subjected to mass spectrometric identification.
Fig.11: Presence of Tic110 in the Tap-tagged Toc33 complex.
Presence of Tic110 in protein eluates was checked by Western blot using antibodies raised against atTic110 on untransformed Col0 and ppi1 plants (A), plants expressing NTAPi-Toc33G (B) and NTAPi-Toc33 (C). (1: starting material, 2: unbound protein, 3: TEV eluates, 4: remaining proteins left on IgG beads). Protein sizes are indicated in kDa. (L: protein molecular weight marker).

Fig.12: Presence of actin in the Tap-tagged Toc33 complex.
Presence of actin in protein eluates was checked by Western blot using antibodies raised against actin on untransformed Col0 and ppi1 plants (A), plants expressing NTAPi-Toc33 (B) and NTAPi-Toc33G (C) and on plants expressing NTAPi-Toc33G that were extracted without TritonX-100 (T-) (D). (1: starting material, 2: unbound proteins, 3: TEV eluates, 4: remaining proteins left on IgG beads). Protein sizes are indicated in kDa. (L: protein molecular weight marker).
atToc33, atTic110, a protein similar to psTic62, heat shock proteins and chaperones were identified as well as actin, tubulin, elongation factors and calmodulin binding proteins (Table 1).

For confirmation of the mass spectrometry results, immunoblots on existing membranes were performed with antibodies raised against atTic110 and actin (Fig. 11 and 12) (these membranes correspond to the small scale TAP tag experiments presented in Fig. 9).

Immunoblotting using α-atTic110 (Fig. 11) failed to detect atTic110 in eluates of control experiment using extracts from Col0 and ppi1 plants (Fig. 11A). A very faint band corresponding to atTic110 was present in NTAPi-Toc33G (Fig. 11B) and NTAPi-Toc33 (Fig. 11C) eluates.

Immunoblotting using α-actin (Fig. 12) showed that actin was not present in the eluates of control experiment using extracts from Col0 and ppi1 plants (Fig. 12A). On the other hand, actin co-purified with NTAPi-Toc33G (Fig. 12B) and NTAPi-Toc33 (Fig. 12C). In plants expressing either NTAPi-Toc33 or NTAPi-Toc33G that were extracted without detergent prior to IgG affinity chromatography (Fig. 12D), actin was not present in the eluates. However, actin could be detected after extraction of the IgG beads with a SDS-containing buffer.

6 Discussion

6.1 Interaction of Toc33 with Toc159 and Toc75 depends on its membrane-association

6.1.1 Complementation of the ppi1 phenotype and chlorophyll accumulation defect by NTAPi-Toc33

ppi1 NTAPi-Toc33G plants were much paler than ppi1 NTAPi-Toc33 plants. Chlorophyll measurements (Fig. 3) showed that chlorophyll accumulation was defective and as low in ppi1 NTAPi-Toc33G plants as it is in ppi1 plants (Jarvis et al., 1998). This result suggests that Toc33 needs to be inserted to chloroplast outer membrane to play its role during import, and therefore only the full length Toc33 having the C-terminal transmembrane sequence can complement the ppi1 phenotype.

We therefore tested whether NTAPi-Toc33 and NTAPi-Toc33G were inserted into the chloroplast outer membrane. We could not visualise the insertion of the fusion proteins directly by immunolocalisation (Fig. 4). This could be explained by the fact that much of NTAPi-Toc33 and NTAPi-Toc33G were soluble and that we could not distinguish the membrane-bound fusion proteins from the cytosolic ones by immunofluorescence. However, after ultracentrifugation of total plant protein extracts, we could discriminate
soluble and membrane-associated fractions of NTAPI-Toc33 and NTAPI-Toc33G (Fig. 5). In *ppi1* plants, NTAPI-Toc33 was only in the membrane fraction. In Col0 plants however, a smaller fraction of NTAPI-Toc33 appeared soluble. As NTAPI-Toc33 was expressed under the 35S promoter, it seems possible that Toc33 insertion sites were saturated and that not all of the molecules were membrane-inserted. Particularly in Col0 plants where endogenous atToc33 is present, such insertion “sites” for Toc33 may be more easily saturated. In the same experiment, part of the NTAPI-Toc33G fusion protein was present in the pellet fraction; this may be due to the over expression of the fusion protein that leads to assembly of a non-physiological complex or alternatively NTAPI-Toc33G may have aggregated.

6.1.2 Differences in Toc proteins association with NTAPI-Toc33 and NTAPI-Toc33G

Our TAP tag purification experiments suggest that the C-terminal transmembrane region of Toc33 is required for interaction with the other members of the Toc core complex. Indeed, by Western blotting, we showed that atToc159 and atToc75 co-purified with NTAPI-Toc33. When NTAPI-Toc33G, lacking the C-terminal transmembrane region, was isolated, neither Toc159 nor Toc75 were detected. The notable exception was Toc159, that was detected with Col0 NTAPI-Toc33G glycine eluates. This suggests that a small part of Toc159 may have been trapped in the cytosol by the soluble G-domain of Toc33. atToc132 also co-purified with NTAPI-Toc33. Since not Toc33 but Toc34 was proposed to interact preferentially with atToc132 and/or atToc120 (Ivanova *et al.*, 2004), it will be interesting to test for the presence of atToc120 or atToc34 in NTAPI-Toc33 eluates.

The crystal structure of pea Toc34 (Sun *et al.*, 2002) and its homologue atToc33 (Koenig *et al.*, 2008), show homodimers of the two proteins. We therefore expected the presence of endogenous Toc33 within the NTAPI-Toc33 purified complex. Unfortunately, we were unable to detect endogenous Toc33 using our α-Toc33 antibodies. In the future, the homodimerisation hypothesis could be addressed by co-expressing two different tagged proteins *in planta*; for example Toc33-TAP tagged and Toc33-His tagged proteins.
6.2 Hypothetic partners revealed by mass spectrometry

6.2.1 Tic110 and chaperones

Mass spectrometry analysis revealed the presence of atTic110 peptides (Table1). We then confirmed by Western blot (Fig.11) the presence of atTic110 in the eluted TAP-tagged Toc33 protein complexes. Signals were faint but atTic110 was detected in both NTAPi-Toc33G and NTAPi-Toc33 TEV eluates. However, atToc75 and atToc75 were not detected by Western blot in NTAPi-Toc33G eluates, which raise the question whether atTic110 is a contaminant, or an interacting partner. Mass spectrometry also revealed the presence of several chaperones in our TEV eluates: cpn60A (Hsp60), Hsp70, Hsp93 and co-chaperones DnaJ proteins (Table1). Chaperones are involved during protein import: Hsp70, for example, is involved in chloroplast import (Schnell et al., 1994) and plays a role in the guidance complex (May and Soll, 2000). Hsp70-protein interacts with a recently identified the DnaJ-like protein, Toc12, during translocation of preproteins (Becker et al., 2004a). Moreover, it was demonstrated that Hsp70 binds to the transit peptides of preproteins (Ivey et al., 2000; Rial et al., 2000).

Cpn60 helps the folding process of newly imported proteins (Kouranov et al., 1998; Inaba et al., 2003). Tic110 is known to recruit chaperones Hsp93 and Cpn60 with the help of Tic40 (Kessler and Blobel, 1996; Kessler and Schnell, 2006; Kovacheva et al., 2005, Kovacheva et al., 2007) to complete precursor translocation across the inner membrane and aid in subsequent protein folding.

The presence of Tic110 in NTAPi-Toc33 complexes suggests a role for Toc33 in establishing contacts between the Toc and Tic complexes.

6.2.2 Actin

Mass spectrometry analysis revealed the presence of actin in the eluted Tap-Tagged Toc33 protein complexes which was confirmed by Western blot (Fig.12). Actin was detected in both NTAPi-Toc33G and NTAPi-Toc33 TEV eluates. However, when NTAPi-Toc33G proteins were extracted in a Triton-free buffer, actin was not detected, in the TEV eluates.

A possible link between Toc proteins and actin is interesting. Actin takes part in several “chloroplastic processes“. Actin plays a role in chloroplast movement (Kandasamy and Meagher, 1999). Actin interacts with a protein involved in chloroplast movement in response to high or low light exposure: CHUP1 (for chloroplast unusual positioning phenotype) (Oikawa et al., 2003; Schmidt von Braun and Schleiff, 2008). Actin (actin cytoskeleton) is also involved in stromules movement. Stromules are stroma-filled
tubules that extend from the chloroplast surface and permit the exchange of material between interconnected chloroplasts (Gray et al., 2001).

6.2.3 α-tubulin and EF1α proteins

α-tubulin protein peptides were found in our eluates. α- and β-tubulins are known to compose microtubules that are involved in different protein processes including vesicular transport. In plant cells, recent evidence suggests that some detergent-resistant membranes called lipid rafts may play an important role in protein targeting (as introduced in Borner et al., 2005). Proteomics analysis on these lipid rafts in Arabidopsis detected the presence of atToc75 (Borner et al., 2005). Moreover, proteomics analysis on these lipid rafts in tobacco cells detected α- and β-tubulin and 14-3-3 like proteins (Mongrand et al., 2004). We could then imagine that some proteins, including components of the Toc core complex, could be targeted to the chloroplast membrane by these lipid rafts with the help of microtubules.

Elongation factor 1 protein peptides were also detected by mass spectrometry. EF1α is known to be an abundant protein that interacts with the cytoskeleton by binding and bundling actin filaments and microtubules, and it participates in actin polymerization (Murray et al., 1996). EF1α was also found by proteomics in Arabidopsis in lipid rafts (Borner et al., 2005).

6.2.4 Calmodulin

Mass spectrometry also identified two calmodulin isoforms. Recent data showed that calcium and calmodulin inhibitors affected protein import, which suggests that the import of chloroplast proteins is regulated by calcium and calmodulin (Chigri et al., 2005). Tic32 was found to interact with calmodulin (Chigri et al., 2006). As the calmodulin isoforms involved in protein import have not been identified yet, the two proteins identified in our complexes may be interesting candidates.

6.2.5 Missing interaction partners

According to the Western blots performed on our TAP tag purification assays, we were strongly expecting to identify peptides of atToc159, atToc75 and atToc132. Surprisingly, apart from Toc33 no component of the Toc core complex was detected by mass spectrometry in our TEV eluates. We may need to scale up the experiment to
detect these proteins by mass spectrometry. On one hand, we could imagine that atToc33 may have role in the formation of Toc-Tic supercomplexes and indirectly in the recruitment of atTic110 and chaperones. On the other hand, due to the high expression level of the fusion proteins, many NTAPi-Toc33 and NTAPi-Toc33G molecules may have aggregated which could prevent the association to atToc159, atToc75 or atToc132 and could explain why these Toc proteins could not be identified by mass spectrometry.

7 Future work to improve isolation and purification of Toc proteins associated complexes

Our first data show that the TAP tag method is suitable to purify Toc33. We will have now to improve this system to find interacting partners, primary by scaling up our experiment and by testing the second purification step on calmodulin-beads. The use of crosslinker chemicals could also be interesting to enlarge Toc33 protein interaction network.

Overexpression of a protein of interest can be problematic as protein-protein interactions are known to take place in a crowded cellular environment and heterogeneous; interaction pairs may exist at low relative concentrations.

Several modelling studies predict that they are many more risks for unspecific interactions compared with specific interaction possibilities (Deeds et al., 2007). For example, in the TAP tag method paper of Puig et al., 2001, the authors cite that overexpression of protein induces association with non natural partners such as heat shock proteins. The TAP tag method is successfully working with proteins expressed at their natural levels (Rigaut et al., 1999), and the next step of this project is to repeat work with plants expressing NTAPi-Toc33/Toc33G under endogenous promoter and terminator.

We also have to consider the fact that proteins may interact with the N-terminal of Toc33 proteins. These studies will have to be carried out with plants expressing Toc33 fused to the TAP-tag in C-terminal (Toc33G-CTAPI and Toc33-CTAPI), although using a C-terminal tag for Toc33 will certainly prevent the protein from being inserted to the membrane.
Materials and Methods

1 Material

1.1 Plants

In Chapter 1, wild type (wt) or Col2 Arabidopsis plant always refers to Arabidopsis thaliana (L.) Heynh. var. Columbia 2. Tobacco plants refer to Nicotiana tabacum cv Petit Havana SR1. In Chapter 2, Wild type (wt) or Col0 Arabidopsis plant always refers to Arabidopsis thaliana (L.) Heynh. var. Columbia O. ppi1 plants are a gift from Dr. Paul Jarvis, University of Leicester. UK.

1.2 Bacteria and yeast strains

Escherichia coli DH5α and Agrobacterium tumefaciens strain GV3101 are from Invitrogen. Yeast strain used in split-ubiquitin assays are JD53 (Dohmen et al. 1995).

1.3 Chemicals

Unless stated otherwise, the chemicals were purchased from Fluka Chemie GmbH.

1.4 Oligonucleotides and Plasmids

Oligonucleotides were synthesized at Microsynth (Switzerland). Plasmids containing yeast split-ubiquitin cassettes STE14-Cub-RURA3, PEX11-Cub-RURA3 and Nub-PEX11 were a gift from prof. Nils Johnsson (University of Münster, Germany). Nub (amino acids 1-37 of ubiquitin) fusions are expressed from a pRS314 plasmid under control of the \( P_{CUP1} \) promoter, Cub-RUra3p (amino acids 37-76 of ubiquitin) fusions from a pRS313 vector under control of the \( P_{MET17} \)-promoter (Eckert and Johnsson, 2003).
pCL60 (described in Bauer et al., 2000) is a pBluescript SK-(Stratagene) derivative with a CaMV 35S promoter and a nopaline synthase (nos) terminator cassette containing the coding sequence for enhanced green fluorescent protein (GFP).

The NTAPI cassette was cloned by Sibylle Infanger into the pCHF8 vector. pCHF8 was kindly provided by Dr. C. Frankhauser (Center for Integrative Genomics, University of Lausanne, Switzerland). pCHF8 is a pCAMBIA3300 derivative with the cauliflower mosaic virus (CaMV) 35S promoter and the RuBisCO small subunit terminator flanking the multiple cloning site.

1.5 Antibodies

Polyclonal antibodies specific to atToc159, atToc132, atToc75 have been described (Bauer et al., 2000; Hiltbrunner et al., 2001b; Ivanova et al., 2004). Anti-Toc159G antibodies were obtained and purified by Dr Birgit Agne. Anti-Toc159A antibodies were done by Dr Birgit Agne and purified by Anouk Sester (master thesis 2008). Antibodies raised against atTic110 were obtained from Dr. Mayte Alvarez-Huerta (University of Utrecht, The Netherlands). Antibodies specific to atToc33 were done and purified with the help of Sylvain Bischof (master thesis 2006).

Monoclonal actin antibodies (clone 10B3 recognizing all eight Arabidopsis actins isoforms (ACT1, 2, 3, 4, 7, 8, 11 and 12)) and IgG antibodies (Human IgG Purified Immunoglobulin Reagent Grade) were purchased at Sigma-Aldrich.

α-HA monoclonal antibodies were bought from Roche, α-Myc monoclonal antibodies from Santa Cruz Biotechnology, and antibodies against phosphoribulokinase (PRK) were a gift from Dr. Pia Stieger, University of Neuchâtel.

2 Methods

2.1 Methods used for cloning

2.1.1 PCR, restriction digest and ligation

Standard protocols were used for cloning (Sambrook and Russell, 2001). DNA fragments were amplified using a proof-reading DNA polymerase (Pwo, Roche) and oligonucleotides including appropriate restriction sites. Pre-existing plasmid-constructs were used as templates for PCR. PCR products and plasmids were digested (restriction enzymes from New England Biolabs or Promega) and purified from agarose gels using
the QIAquick kit (Quiagen). Vectors were dephosphorylated using shrimp alkaline phosphatase (Roche) according to the manufacturer’s recommendations. T4 DNA ligase (New England Biolabs) was used to ligate vectors and inserts as described (Sambrook and Russell, 2001). PCR reactions and purifications steps were analysed by agarose gel electrophoresis according to standard protocols.

2.1.2 Transformation of chemically competent E. coli cells and plasmid isolation

Ligation reactions were subsequently transformed by heat shock into competent E. coli DH5α cells (Sambrook and Russell, 2001). Competent E. coli cells were prepared as described (Inoue et al., 1990). Selection was done on LB medium [25g/l LB Broth Miller (Becton Dickinson Diagnostic Systems), 1.2% (w/v) Agar Bacteriological Grade (ICN Biomedicals)] supplemented with appropriate antibiotics. Clones were selected by PCR and restriction digestion and DNA sequences were verified by sequencing (Microsynth).

For plasmid isolation, the GenElute Plasmid Miniprep Kit (Sigma) was used according to the suppliers’ instructions.

2.2 DNA constructs

2.2.1 Yeast split ubiquitin cloning

The yeast split-ubiquitin constructs were derived from the STE14-Cub-RURA3 (Wittke et al., 1999), PEX11-Cub-RURA3 and Nub-PEX11 constructs (Eckert and Johnsson 2003). Two hemagglutinin epitopes (HA) were added to the Cub constructs by annealing the primers 5’-TCG ACC TAC CCA TAC GAC GTA CCA GAT TAC GCT GCT TAC CCA TAC GAC GTA CCA GAT TAC GCT-3’ and 5’-TCG AAG CGT AAT CTG GTA CGT CGT ATG GGT AAG CGT AAT CTG GTA CGT CGT ATG GGT AGG-3’ and ligation into the unique SalI restriction site in front of the Cub coding sequence.

The coding sequence of the G domain of Toc33 (amino acids 1-265) was amplified from existing plasmid pet21d-Toc33H6sol using a forward primer containing a Clal restriction site 5’-CCA TCG ATC CAT GGG GTC TCT CG-3’ and a reverse primer including a SalI site 5’-CAT ATG GTC GAC CCT ATC TTT CCT TTA TCA TC-3’ and cloned into the Clal/SalI digested STE14-Cub-RURA3 construct (Wittke et al., 1999). The coding sequence of the G domain of Toc159 (amino acids 728-1093) was amplified from pet21d-Toc159G using the following forward primer containing the coding sequence for a single Myc epitope tag and a BamHI site 5’-CCC GGG ATC...
The coding sequence of preats1B (AT5G38430.1) was amplified from pET21d-preats1B using forward primer 5'-CAG GGA TCC ATC GAT GAT ATC CTC TAT CCT CTA GAG TCG ACC ATG CAG ATC TTC GCT C-3' containing a \textit{ClaI} restriction site, and reverse primer 5'-CC GCT CGA GCC AGC ATC AGT GAA GCT TC-3' that contains a \textit{XhoI} site. The coding sequence of ats1B (amino acids: 55-181) was amplified from pET21d-preats1B using forward primer 5'-G TAG GAT CCA TCG ATG AAG GTG TGG CCA C-3' that contains a \textit{ClaI} restriction site and the same reverse primer used to amplify preats1B. The PCR products were then digested by \textit{ClaI}/\textit{XhoI} and cloned into the \textit{ClaI} \textit{SalI} digested \textit{STE14-Cub-RURA3} construct.

Mutations of Nub Ile13 to Gly (Nub\textsuperscript{G}) or Ile13 to Ala (Nub\textsuperscript{A}) were done by site-directed mutagenesis according to the manufacturer’s manual (Quick change, Invitrogen). For Nub\textsuperscript{G} primers used were as follow: forward 5'-G ACT TTG ACC GGT AAA ACC GG A ACA TTG GAA GTT GAA TCT TC-3' and reverse primer containing a \textit{NcoI} site 5'-TCA TGT CAT GAC ACC ACC GCG GAG ACG G-3'. For Nub\textsuperscript{A} we used: forward 5'-G ACT TTG ACC GGT AAA ACC GC A ACA TTG GAA GTT GAA TCT TC-3' and reverse 5'-GA AGA TTC AAC TTC CAA TGT T-3'.

2.2.2 Plant split-ubiquitin constructs

The plant split-ubiquitin constructs were designed with the coding sequence of plant ubiquitin atUBQ11 (AT4G05050.1) (Callis \textit{et al.} 1995). The sequence corresponding to the first 37 amino acids (Nub) was amplified using as a forward primer 5'-CGG GAT CCT CTA GAG TCG ACC ATG CAG ATC TTC GCT C-3' including a \textit{BamHI} site and a reverse primer containing a \textit{NcoI} site 5'-TCA TGT CAT GAC ACC ACC GCG GAG ACG G-3'. A plasmid (BUGUS) containing the atUBQ11 coding sequence, provided by Prof. Richard Vierstra (University of Wisconsin-Madison) served as template. The resulting PCR fragment was ligated into the vector pCL60 cut by \textit{BamHI} and \textit{NcoI} yielding pCL60-Nub. The sequence corresponding to the amino acids 35-76 of UBQ11 (Cub) was amplified using the forward primer 5'-CAT GCC ATG GGA TAC CCA TAC GAC GTA CCA GAT TAC GCT GGC ATT CCT CCG GAC C-3' including a \textit{NcoI} site and the coding sequence for a single HA tag and the reverse primer 5'-TCA TGT CAT GAC ACC ACC GCG GAG ACG G-3' containing a \textit{BspHI} site. The PCR product was ligated into pCL60 vector cut by \textit{NcoI} yielding pCL60-Cub.
The following primers 5’-GTA CTC ATG AAG GAG CAG AAG CTG ATC-3’ (forward), 5’-CTC AAG ACC CGT TTA GAGG-3’ were used to amplify Toc159<sub>728-1093</sub> (Toc159G) with the two-hybrid construct pGBK7-Toe159G as DNA template. The amplified DNA was then cloned *NcoI* and *NolI* into pCL60-Nub. The complete sequence of Toc33 or Toc33G (Toc33<sub>1-265</sub>) were amplified with the forward primer 5’-TGG GCC ATG GGG TCT CTC GTT CGT-3’ and the reverse primers 5’-TGA ACT CAT GAG TGA TTT CTT CAA C-3’ or 5’-TGA ACT CAT GAG CTT TCC TTT ATC ATC-3’, respectively. Ligation was done in pCL60-Cub vector cut by *NcoI*. The coding sequence of atPex1<sub>1-2</sub> (atPEX1e, AT3G61070.1) was amplified by the forward primer 5’-CAT GCC ATG GCA ACT ACA CTA GAT TTG ACC-3’ containing a *NcoI* site and the reverse primer 5’-CTA TAG CGG CCG CTC ATG ATT TCT TCA AC-3’ including a *NolI* site. The product was ligated into pCL60-Nub cut by *NcoI* and *NolI*. To clone into pCL60-Cub cut by *NcoI*, atPex11-2 was amplified with the same forward primer as above and reverse 5’-TGA ACT CAT GAG TGA TTT CTT CAA C-3’ including a *BspHII* site. The template plasmid DNA pGEM-Teasy-PEX11.2 was kindly provided by the group of Alison Baker (University of Leeds, UK).

Mutation of Nub Ile13 to Gly (Nub<sup>G</sup>) was done by site-directed mutagenesis according to the manufacturer’s manual (Quick change, Invitrogen) using forward primer: 5’-CC GGA AAG ACC GG C ACT CTT GAA GTT GAG AGT TCC GAC ACC-3’ and reverse primer: 5’-GGT GTC GGA ACT CTC AAC TTC AAG AGT GGG GGT CTT TCC GG-3’.

### 2.3 Stable transformation of *Arabidopsis* with TAP tag constructs.

cDNAs of Toc33 and Toc33G were amplified by PCR from pCL60-Toe33-Cub and pCL60-Toe33G-Cub (see split ubiquitin chapter) using as a forward primer 5’-GAA GAG GCC TTT ATG GGG TCT CTC GTT-3’ including *StuI* site and reverse primers 5’-G CTC TAG AGC TTA AAG TGG CTT TCC TTT ATC-3’ and 5’-G CTC TAG AGC TTA AAG TGG CTT TCC TTT ATC-3’ containing *XbaI* site to amplify Toc33 and Toc33G respectively. Toc33 and Toc33G amplified fragments were digested by *StuI/XbaI* and ligated in the corresponding sites of pCHF8-NTAPi leading to pCHF8-NTAPi-Toe33 and pCHF8-NTAPi-Toe33G.

Electrocompetent GV3101 cells were prepared as described (Weigel and Glazebrook, 2002). 30ng of plasmid DNA were used for electroporation in a MicroPulser (BioRad) device, according to the manufacturer’s instructions. Transformants were selected on LB media supplemented with 1.2% (w/v) agar (bacteriological grade, ICN Biomedicals) and appropriate antibiotics.
A. thaliana plants were transformed using the floral dip method as described (Clough and Bent, 1998). Transformants were selected on plates or on soil by phosphinothricin (see below).

2.4 Plant growth

2.4.1 Growing plants in vitro

Seeds were surface-sterilised by liquid or vapour-phase methods as described (Clough and Bent, 1998). Col0 and ppi seedlings were plated on 0.5x Murashige and Skoog medium (Duchefa) containing 0.8% Phyto Agar (Duchefa) and left for 2 days at 4°C in the dark. They were then grown under short-day conditions (8h light (120μmol m^{-2} s^{-2}) 16h dark, 20°C, 70% relative humidity). NTAPi-Toc33/Toc33G transformed seedlings were selected by adding in the Murashige and Skoog medium 30 μg/ml phosphinothricin (BASTA; Duchefa).

Nicotiana tabacum cv Petit Havana SR1 were grown on 1x Murashige and Skoog medium containing 0.8% Phyto Agar under long-day conditions (16h light 120μmol m^{-2} s^{-2}, 8h dark, 23°C, 60% relative humidity).

2.4.2 Growing Arabidopsis on soil

Seeds were planted on soil (Ricoter). Germination was synchronised as above. Plants were grown under short day conditions (8h light (120μmol m^{-2} s^{-2}) 16h dark, 20°C, 70% relative humidity) or long-day conditions (16 h light (120μmol m^{-2} s^{-2}), 8 h dark), 20°C, 70% relative humidity) to get seeds. NTAPi-Toc33/Toc33G transformed plants were selected by spraying on seedlings 30 μg/mL phosphinothricin (BASTA; Duchefa) every 3 days.

2.5 Yeast split-ubiquitin assay

Yeast growth was performed as described (Johnsson and Varshavsky, 1994). Total protein extracts were prepared according to Kiel et al., 2005.
2.6 Protoplasts transformation

Protoplasts were transiently transformed using the polyethylene glycol method according to (Jin et al. 2001) with 4 weeks-old A. thaliana or 6 weeks-old N. tabacum leaves.

Fluorescence in transformed protoplasts was monitored 24 to 48h after transformation using a Leica TCS 4D microscope. GFP was detected with the FITC (488 nm) laser line and TRITC (568 nm) was used for chlorophyll autofluorescence.

2.7 Chlorophyll measurements

Measurements were done according to Sundberg et al. 1997 using Col0, ppi1, Col0/ppi NTAPI-Toc33/Toc33G seedlings. Experiment was repeated 10 times for each different seedlings.

2.8 Plant genomic DNA extraction and PCR analysis

Leaf tissue corresponding to 0.5cm² was ground in 0.2M Tris/HCl pH9, 0.4M LiCl, 25mM EDTA, 1% (w/v) SDS. The homogenate was cleared by centrifugation (5min, 16000xg, 4°C), supplemented with an equal volume of isopropanol and centrifuged again (45min, 16000xg, rt). The supernatant was discarded and the DNA pellet was washed with 80% (v/v) ethanol. Finally, DNA was dried and resuspended in 50µL TE (10mM Tris/HCl pH8, 1mM EDTA).

PCR reactions contained 2µL DNA solution, 0.5U GoTaq DNA polymerase (Promega), as well as 0.2mM dNTPs (Eurobio) in a total volume of 25µL. Toc33 gene presence was monitored using forward 5'-GGT CTC TCG TTC GTG AAT GG-3' and reverse 5'-TTG GAA AGG ACT GAC ACG GA-3' primers. ppi1 genotype was checked using same forward primer as above and the following T-DNA left border reverse primer 5'-ATA ACG CTG CGG ACA TCT AC-3'. The PCR program was as follows: 5min, 94°C, 35x (30s, 94°C; 30s, 55°C; 30s, 72°C).
2.9 Plant proteins extraction and Western blot analysis

Transiently transformed protoplasts were centrifuged for 1 min at 100xg, or proteins were isolated from Arabidopsis leaves. Total proteins were extracted according to Rensink (100mM NaCl, 50mM TrisHCl pH7.5, 1mM PMSF, 0.5% v/v TritonX-100) (Rensink et al.,1998). To avoid proteolytic degradation, 0.5% (v/v) protease inhibitor cocktail for plant cell extracts (Sigma P9599) was added to the extraction buffer. Proteins were concentrated by chloroform-methanol precipitation (Wessel and Flügge, 1984) and resuspended in sample buffer [50mM Tris/HCl pH6.8, 0.1M DTT, 2% (w/v) SDS, 0.1% (w/v) bromophenol blue, 10% (v/v) glycerol] and heated 10min at 65°C. Protein concentration was determined by the Bradford assay (Bradford 1976) and reagent purchased from BioRad using bovine serum albumine (BSA) as the standard.

Proteins were separated by SDS-PAGE and blotted onto Protran R nitrocellulose membrane (Schleicher & Schuell) using the Mini-PROTEAN System (BioRad). Western blot membranes were stained with Amido Black (Naphthol Blue Black) as described (Sambrook and Russell, 2001) and scanned.

To block unspecific binding of antibodies, membranes were incubated 1h or o/n in blocking buffer [PBS (140mM NaCl, 2.7mM KCl, 10mM Na2HPO4, 1.8mM KH2PO4, pH7.4), TBS (25mM Tris/HCl pH7.4, 140mM NaCl, 2.7mM KCl) containing 5% (w/v) skim milk powder]. Membranes were then incubated for 1h or longer with adequate antibody dilutions in blocking buffer. (see below) and washed extensively with PBS, TBS or TBS-tween (TBST; TBS, 0.05% (v/v) Tween R 20). To reveal primary antibodies, membranes were incubated 1h with a 3’000x dilution of horseradish peroxidase-coupled goat anti-rabbit or goat anti-mouse IgG (Bio-Rad Laboratories) in TBS-T buffer containing 1% (w/v) skim milk powder. After washing with PBS/TBS/TBS-T buffer, signals were detected by enhanced chemiluminescence.

The TAP tag experiments membranes were incubated 1 min in [0.1M Tris/HCl pH8.5, 1.25mM 3-aminophthalhydrazide (luminol), 0.2mM p-coumaric acid, 0.009% (v/v) H2O2] and exposed to high performance chemiluminescence films (Amersham Biosciences).

Split-ubiquitin experiments blots were developed using ECL and high performance films (GE Healthcare). Chemiluminescence signals were quantified using ImageJ (http://rsb.info.nih.gov/ij/).

Antibody dilutions for immunoblotting, Incubation buffers and approximate incubation times for immunoblotting are indicated. Unless stated otherwise, antibodies were produced into rabbit organisms.

α-atToc159A (1/2000) PBS 5% Milk powder, 1h rt
α-atToc159G (1/200) PBS with no milk powder, o/n 4°C
α-atToc132 (1/2000) PBS 3% Milk powder, o/n 4°C
\(\alpha\)-atToc75 (1/3000) PBS 3% Milk powder, 2h rt
\(\alpha\)-atToc33 (1/200) PBS 5% Milk powder, o/n 4°C
\(\alpha\)-atTic110 (1/3000) PBS 3% Milk powder, 2h rt
\(\alpha\)-actin (1/2000) PBS 1% Milk powder, 1h rt, mouse
\(\alpha\)-IgG (1/5000) PBS 5% Milk powder, 1-2h rt, human
\(\alpha\)-PRK (1/500) PBS 5% Milk powder, 1h rt
\(\alpha\)-HA (1/500 or 1/1000) TBS 1% Milk powder, 2h rt or o/n 4°C, mouse
\(\alpha\)-Myc (1/5000) TBS 1% Milk powder, 1h rt, mouse

[PBS (140mM NaCl, 2.7mM KCl, 10mM Na2HPO4, 1.8mM KH2PO4, pH7.4), TBS (25mM Tris/HCl pH7.4, 140mM NaCl, 2.7mM KCl), TBST (TBS, 0.05% (v/v) Tween R 20)].

2.10 Separation of soluble and insoluble proteins

Transformed protoplasts were collected by centrifugation at 100xg for 1min and resuspended in lysis buffer (20mM Tris-HCl, pH 7.5, 50mM NaCl, 1mM DTT, 2mM MgCl2, 0.5% w/v inhibitor cocktail for plant cell extracts) followed by freezing and thawing. The lysate was centrifuged at 100000xg for 1h at 4°C; the resulting supernatant was considered total soluble protein. Soluble protein was concentrated by chloroform-methanol precipitation. The pellet was resuspended with 50 mM Tris-HCl, pH7.5.

2.11 Immunolocalization on protoplasts

Protoplasts were prepared as described (Jin et al., 2001), reducing cellulase and macerozyme (Serva) concentrations to 1% and 0.25% (w/v) respectively.

Immunolocalization in isolated protoplasts was performed as described (Matsui et al. 1995) Polysine slides (BDH Laboratories) were used. Protoplasts were fixed with 4% (v/v) formaldehyde in W5 buffer [154mM NaCl, 125mM CaCl2, 5mM KCl, 5mM glucose, 1.5mM MES pH5.6] and permeabilized with 0.5% (v/v) NP-40 in W5. Primary antibodies were diluted 1:100 (v/v, anti-IgG) and to 1/20 (v/v, anti-atToc33). Secondary antibody was diluted at 1/100. Slides were mounted in SlowFade solution (Molecular Probes) and fluorescence of the fluorescein-coupled secondary antibody (Pierce) was monitored by confocal scanning microscopy using the FITC (488 nm) laser line from a LEICA TCS SP5 microscope.
2.12 TAP tag purification

Purification of TAP tag complexes from plants expressing tagged Toc33 was done as described (Rohila et al., 2004, Rubio et al., 2005, Witte et al., 2005) using commercial IgG-sepharose beads (Amersham). Plant leaves were ground with liquid Nitrogen. Total plant proteins were extracted as described in section 2.9 (according to Rensink et al., 1998) using or not using 0.5% v/v Triton X-100 in the extraction buffer depending on the experiment. Following this step, proteins were centrifuged at 10’000g for 30min. When experiment was done starting with 10g of plants, proteins were ultracentrifuged at 100’000g for 1h. This centrifugation step is necessary to separate solubilized and non-solubilized material. Protein concentration was determined by the Bradford assay. Proteins (e.g. centrifuged supernatant) were incubated to beads for 2h30min at 4ºC on a rotating wheel in 1.5mL microtubes. Proteins were eluted with 0.1M Glycine pH3 (acidic elution) or with TEV protease (Promega). For acidic elution, beads were incubated 5min on ice with 100µL 0.1M Glycine pH3. This step was repeated 3 times. TEV elution was processed according to the manufacturer manual, adding 1µM protease inhibitor E64 Roche (Rohila et al., 2004) to the cleavage buffer. Beads were incubated with TEV protease for 2h at 16ºC or o/n at 4ºC. Samples of the different steps of the experiment were concentrated and resuspended as described in section6. Sample buffer was added to the eluted beads to load on SDS gel.

2.13 Protein identification by mass spectrometry

TEV eluates were subjected to SDS-PAGE and proteins were stained with SYPRO Ruby (Invitrogen, according to the manufacturer protocol except that methanol was replaced by ethanol). Visible protein bands were excised and subjected to in-gel tryptic digest. (Fig.11). In gel trypsic digest was perfomed as described (Shevchenko et al., 1996) using for digestion 10ng/µL trypsin (Promega sequence-grade modified porcine trypsin, Cat. #V511A) over night at 30ºC.

Mass spectrometry was performed by Sylvain Bischof, ETH, Zurich:

Mass spectrometry measurements were performed on an LTQ FT-ICR (Thermo Finnigan), coupled with a Probot (LC-Packings/Dionex) autosampler system and the UltiMate HPLC-system (LC-Packings/Dionex). Peptide mixtures were loaded onto laboratory made capillary columns (75µm inner diameter, 8 cm length, packed with Magic C18 AQ beads, 3µm, 100Å (Microm)). Peptides were eluted from the column by an increased acetonitrile concentration in the mobile phase from 5% acetonitrile, 0.2% formic acid to 40% acetonitrile, 0.2% formic acid over 75min, followed by a 10min wash step at 5% acetonitrile, 0.2% formic acid. Peptide ions were detected in a survey scan from 300 to 1’600 amu followed by 3 data-dependent MS/MS scans (isolation
width 2 amu, relative collision energy 35%, dynamic exclusion enabled, repeat count 1, followed by peak exclusion for 2 min).

Interpretation of MS/MS spectra and data filtering:

MS/MS spectra were searched with TurboSequest and PeptideProphet by using the Trans-Proteomic Pipeline (TPP v2.9) against the *Arabidopsis thaliana* TAIR8 protein database (downloaded on December 14th 2007) supplemented with contaminants. The search parameters were: requirement for tryptic ends, one missed cleavage allowed, mass tolerance = +/- 3 Da, variable modification of methionine (M, PSI-MOD name: oxidation, ModAccession: MOD:00412, mono Δ = 15.9949) and static modification of cysteine (C, PSI-MOD name: iodoacetamide derivative, ModAccession: MOD:00397, mono Δ = 57.021464).

For PeptideProphet, the cutoff was set to a minimum probability of 0.9.
Bibliography


Inoue K, Keegstra K (2003.) A polyglycine stretch is necessary for proper targeting of the protein translocation channel precursor to the outer envelope membrane of chloroplasts. Plant J 34, 661–669.


unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res 3, 109-136.


chloroplast import, and accumulation of photosynthetic proteins. Plant Cell 15, 1859-
1871.


Küchler M, Decker S, Hörmann F, Soll J, Heins L (2002). Protein import into
chloroplasts involves redox-regulated proteins. EMBO J 21, 6136-6145.

and cellular approaches for the detection of protein-protein interactions: latest
techniques and current limitations. Plant J 53, 610-635.

essential role in the import of proteins into chloroplasts and chloroplast biogenesis. J
Biol Chem 278, 36794-36805.

system. Brief Funct Genomic Proteomic 1, 230-238.


14, 505-508.

Li X, Gould SJ (2003). The dynamin-like GTPase DLP1 is essential for peroxisome
division and is recruited to peroxisomes in part by PEX11. J Biol Chem 278, 17012-
17020.


Hamacher T, Boles E, von Wirén N, Frommer WB (2003). Homo- and hetero-
oligomerization of ammonium transporter-1 NH4 uniporters. J Biol Chem 278, 45603-
45610.

Ma Y, Kouranov A, LaSala SE, Schnell DJ (1996). Two components of the chloroplast
protein import apparatus, IAP86 and IAP75, interact with the transit sequence during
the recognition and translocation of precursor proteins at the outer envelope. J Cell Biol
134, 315-327.


Smith MD, Hiltbrunner A, Kessler F, Schnell DJ (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.


Appendix

Mass spectrometry results

Mass spectrometry results for each band cut (Fig.10). Proteins identified with less than 3 peptides were not considered, as well as protein peptides presents in both control plants (Col0 and ppi1) and NTAPi-Toc33G/Toc33. In this work, a selection of these proteins was made (Table 1) as we discarded chloroplast metabolism proteins, plastid-lipid associated protein and thylakoids components.

Bands cuts between 200 and 80kDa

<table>
<thead>
<tr>
<th>Col0  NTAPi Toc33</th>
<th>accession number</th>
<th>percent coverage</th>
<th>number of peptides</th>
<th>percent share</th>
</tr>
</thead>
<tbody>
<tr>
<td>atToc33</td>
<td>AT1G02280</td>
<td>24</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Actin</td>
<td>AT5G09810</td>
<td>17</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>tubulin TU A4</td>
<td>AT1G04820</td>
<td>8</td>
<td>3</td>
<td>0,8</td>
</tr>
<tr>
<td>EF1</td>
<td>AT1G07920</td>
<td>10</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>ATSCO1 translation EF</td>
<td>AT1G62750</td>
<td>10</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>CAB</td>
<td>AT1G29910</td>
<td>4</td>
<td>3</td>
<td>0,55</td>
</tr>
<tr>
<td>PATL1 (Patellin 1); transporter</td>
<td>AT1G72150</td>
<td>26,5</td>
<td>13</td>
<td>3,5</td>
</tr>
</tbody>
</table>

Bands between 90 and 66 kDa

<table>
<thead>
<tr>
<th>ppi1 NTAPi Toc33</th>
<th>accession number</th>
<th>percent coverage</th>
<th>number of peptides</th>
<th>percent share</th>
</tr>
</thead>
<tbody>
<tr>
<td>atToc33</td>
<td>AT1G02280</td>
<td>8</td>
<td>3</td>
<td>0,42</td>
</tr>
<tr>
<td>atTic110</td>
<td>AT1G06950</td>
<td>5</td>
<td>3</td>
<td>0,67</td>
</tr>
<tr>
<td>actin</td>
<td>AT5G09810</td>
<td>11</td>
<td>3</td>
<td>0,5</td>
</tr>
<tr>
<td>EF1</td>
<td>AT1G07920</td>
<td>7,8</td>
<td>3</td>
<td>0,62</td>
</tr>
<tr>
<td>Arabidopsis Rab GTPase homolog E1b; translation elongation factor</td>
<td>AT4G20360</td>
<td>12</td>
<td>5</td>
<td>0,9</td>
</tr>
<tr>
<td>myrosinase-associated protein</td>
<td>AT1G54010</td>
<td>10,4</td>
<td>3</td>
<td>0,63</td>
</tr>
<tr>
<td>tubulin</td>
<td>AT1G04820</td>
<td>4</td>
<td>4</td>
<td>0,6</td>
</tr>
<tr>
<td>ESM1 carboxylic ester hydrolase</td>
<td>AT3G14210</td>
<td>17</td>
<td>8</td>
<td>1,9</td>
</tr>
<tr>
<td>Photosystem II subunit O-2</td>
<td>AT3G50820</td>
<td>8,5</td>
<td>3</td>
<td>0,63</td>
</tr>
<tr>
<td>ATMDAR1 (NaDH)</td>
<td>AT3G52880</td>
<td>14</td>
<td>5</td>
<td>0,9</td>
</tr>
<tr>
<td>fructose-1,6-bisphosphatase</td>
<td>AT3G54050</td>
<td>10,6</td>
<td>3</td>
<td>0,66</td>
</tr>
<tr>
<td>Glutamine synthtase 2</td>
<td>AT5G35630</td>
<td>15</td>
<td>5</td>
<td>0,9</td>
</tr>
<tr>
<td>PetA cytF</td>
<td>ATCG00540</td>
<td>22</td>
<td>6</td>
<td>1,13</td>
</tr>
</tbody>
</table>
### Bands 66kDa

**ppi1 NTAPI Toc33**

<table>
<thead>
<tr>
<th>protein</th>
<th>accession number</th>
<th>percent coverage</th>
<th>number of peptides</th>
<th>percent share</th>
</tr>
</thead>
<tbody>
<tr>
<td>atTic110</td>
<td>AT1G06950</td>
<td>12,6</td>
<td>10</td>
<td>1,5</td>
</tr>
<tr>
<td>Actin</td>
<td>AT5G09810</td>
<td>12</td>
<td>3</td>
<td>0,5</td>
</tr>
<tr>
<td>CPN60A chaperonine</td>
<td>AT2G28000</td>
<td>23</td>
<td>12</td>
<td>1,8</td>
</tr>
<tr>
<td>CPN60B chaperonine</td>
<td>AT1G55490</td>
<td>10,6</td>
<td>5</td>
<td>0,8</td>
</tr>
<tr>
<td>cpHSC70-2 (Heat shock protein 70-7)</td>
<td>AT5G49910</td>
<td>15</td>
<td>10</td>
<td>0,9</td>
</tr>
<tr>
<td>Heat shock protein (HSP) 93-V</td>
<td>AT5G50920</td>
<td>16</td>
<td>12</td>
<td>1,8</td>
</tr>
</tbody>
</table>

### Bands between 66 and 35kDa

**col0 NTAPI Toc33G**

<table>
<thead>
<tr>
<th>protein</th>
<th>accession number</th>
<th>percent coverage</th>
<th>number of peptides</th>
<th>percent share</th>
</tr>
</thead>
<tbody>
<tr>
<td>atToc33</td>
<td>AT1G02280</td>
<td>47</td>
<td>52</td>
<td>11</td>
</tr>
<tr>
<td>atTic110</td>
<td>AT1G06950</td>
<td>19</td>
<td>15</td>
<td>3,1</td>
</tr>
<tr>
<td>Actin</td>
<td>AT5G09810</td>
<td>12</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>porin, putative \ voltage-dependent anion-selective channel protein,</td>
<td>AT5G15090</td>
<td>12</td>
<td>4</td>
<td>0,6</td>
</tr>
<tr>
<td>Elongation factor EF1</td>
<td>AT1G07920</td>
<td>18</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td>CPN60A chaperonine</td>
<td>AT2G28000</td>
<td>20</td>
<td>17</td>
<td>2,5</td>
</tr>
<tr>
<td>CPN60B chaperonine</td>
<td>AT1G55490</td>
<td>16</td>
<td>7</td>
<td>0,9</td>
</tr>
<tr>
<td>cpHSC70-2</td>
<td>AT5G49910</td>
<td>17</td>
<td>20</td>
<td>1,8</td>
</tr>
<tr>
<td>HSP 93-V</td>
<td>AT5G50920</td>
<td>13,5</td>
<td>24</td>
<td>3,34</td>
</tr>
<tr>
<td>FIB (Fibrillin); structural molecule</td>
<td>AT4G04020</td>
<td>12</td>
<td>3</td>
<td>0,4</td>
</tr>
<tr>
<td>Pap fibrillin family</td>
<td>AT3G23400</td>
<td>8</td>
<td>3</td>
<td>0,2</td>
</tr>
<tr>
<td>plastid-lipid associated protein PAP</td>
<td>AT2G35490</td>
<td>9,6</td>
<td>3</td>
<td>0,4</td>
</tr>
<tr>
<td>PAP</td>
<td>AT4G22240</td>
<td>12</td>
<td>5</td>
<td>0,6</td>
</tr>
<tr>
<td>PUR ALPHA-1 (purin-rich alpha 1); nucleic acid binding</td>
<td>AT2G32080</td>
<td>19</td>
<td>5</td>
<td>1,1</td>
</tr>
<tr>
<td>FTSZ1-1 structural molecule</td>
<td>AT5G55280</td>
<td>7</td>
<td>3</td>
<td>0,75</td>
</tr>
<tr>
<td>PORB</td>
<td>AT4G27440</td>
<td>9,7</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>POR C</td>
<td>AT1G03630</td>
<td>6,8</td>
<td>6</td>
<td>0,75</td>
</tr>
<tr>
<td>hydrolase, alpha/beta fold family protein</td>
<td>AT1G52510</td>
<td>11</td>
<td>3</td>
<td>0,75</td>
</tr>
<tr>
<td>ATLFRN1</td>
<td>ATLFRN1 (LEAF FNR 1); NADPH dehydrogenase</td>
<td>oxidoreductase</td>
<td>AT5G66190</td>
<td>14</td>
</tr>
<tr>
<td>Protein</td>
<td>Accession Number</td>
<td>Percent Coverage</td>
<td>Number of Peptides</td>
<td>Percent Share</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>--------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>atToc33</td>
<td>AT1G02280</td>
<td>40</td>
<td>27</td>
<td>5</td>
</tr>
<tr>
<td>atTic110</td>
<td>AT1G06950</td>
<td>12</td>
<td>9</td>
<td>1.4</td>
</tr>
<tr>
<td>Actin</td>
<td>AT5G09810</td>
<td>15</td>
<td>6</td>
<td>1.1</td>
</tr>
<tr>
<td>EF1</td>
<td>AT1G07920</td>
<td>13</td>
<td>4</td>
<td>0.6</td>
</tr>
<tr>
<td>CPN60A chaperonine</td>
<td>AT2G28000</td>
<td>15</td>
<td>8</td>
<td>1.3</td>
</tr>
<tr>
<td>CPN60B chaperonine</td>
<td>AT1G55490</td>
<td>16</td>
<td>8</td>
<td>1.2</td>
</tr>
<tr>
<td>chaperonine 20</td>
<td>AT5G20720</td>
<td>16</td>
<td>3</td>
<td>0.5</td>
</tr>
<tr>
<td>cpHSC70-2</td>
<td>AT5G49910</td>
<td>13</td>
<td>8</td>
<td>1.25</td>
</tr>
<tr>
<td>HSP 93-V</td>
<td>AT5G50920</td>
<td>8</td>
<td>6</td>
<td>0.96</td>
</tr>
<tr>
<td>PAP</td>
<td>AT4G22240</td>
<td>12</td>
<td>3</td>
<td>0.46</td>
</tr>
<tr>
<td>PRK</td>
<td>AT1G32060</td>
<td>6,3</td>
<td>3</td>
<td>0.4</td>
</tr>
<tr>
<td>PORB</td>
<td>AT4G27440</td>
<td>9,5</td>
<td>4</td>
<td>0.74</td>
</tr>
<tr>
<td>POR C</td>
<td>AT1G03630</td>
<td>5</td>
<td>3</td>
<td>0.5</td>
</tr>
<tr>
<td>PSBB (PSII)</td>
<td>ATCG00680</td>
<td>5</td>
<td>3</td>
<td>0.75</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protein</th>
<th>Accession Number</th>
<th>Percent Coverage</th>
<th>Number of Peptides</th>
<th>Percent Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>atToc33</td>
<td>AT1G02280</td>
<td>34</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Actin</td>
<td>AT1G49240</td>
<td>11</td>
<td>5</td>
<td>0.5</td>
</tr>
<tr>
<td>DNAJ heat shock family protein</td>
<td>AT1G07920</td>
<td>10</td>
<td>4</td>
<td>0.8</td>
</tr>
<tr>
<td>PSBB (PSII)</td>
<td>AT1G00680</td>
<td>5</td>
<td>3</td>
<td>0.6</td>
</tr>
<tr>
<td>plastid-lipid associated protein PAP</td>
<td>AT2G35490</td>
<td>9</td>
<td>3</td>
<td>0.6</td>
</tr>
<tr>
<td>oxidoreductase</td>
<td>AT1G23740</td>
<td>11</td>
<td>3</td>
<td>0.5</td>
</tr>
<tr>
<td>AGT (Alanine: glyoxylate aminotransferase)</td>
<td>AT2G13360</td>
<td>12</td>
<td>3</td>
<td>0.5</td>
</tr>
<tr>
<td>ADK1 (Adenosine kinase 1)</td>
<td>AT3G09820</td>
<td>15</td>
<td>4</td>
<td>0.85</td>
</tr>
<tr>
<td>CPN60B (Chaperonin 60 beta)</td>
<td>AT1G55490</td>
<td>6,4</td>
<td>3</td>
<td>0.6</td>
</tr>
<tr>
<td>PTAC16 (Plastid transcriptionally active 18)</td>
<td>AT3G46780</td>
<td>11</td>
<td>4</td>
<td>0.9</td>
</tr>
</tbody>
</table>
### ppi1 NTAPiToc33

<table>
<thead>
<tr>
<th>protein</th>
<th>accession number</th>
<th>percent coverage</th>
<th>number of peptides</th>
<th>percent share</th>
</tr>
</thead>
<tbody>
<tr>
<td>atToc33</td>
<td>AT1G02280</td>
<td>8</td>
<td>3</td>
<td>0,5</td>
</tr>
<tr>
<td>atTic110</td>
<td>AT1G06950</td>
<td>5</td>
<td>3</td>
<td>0,7</td>
</tr>
<tr>
<td>Actin</td>
<td>AT5G09810</td>
<td>11</td>
<td>3</td>
<td>0,5</td>
</tr>
<tr>
<td>EF1</td>
<td>AT1G07920</td>
<td>8</td>
<td>3</td>
<td>0,6</td>
</tr>
<tr>
<td>myrosinase-associated protein</td>
<td>AT1G54010</td>
<td>10</td>
<td>3</td>
<td>0,6</td>
</tr>
<tr>
<td>PSBO2</td>
<td>AT5G66570</td>
<td>8,5</td>
<td>3</td>
<td>0,6</td>
</tr>
<tr>
<td>ESM1 (Epithiospecifier modifier 1) carboxylic ester hydrolase</td>
<td>AT3G14210</td>
<td>17</td>
<td>8</td>
<td>1,7</td>
</tr>
</tbody>
</table>

Bands between 30 and 14 kDa

### col0 NTAPiToc33G

<table>
<thead>
<tr>
<th>protein</th>
<th>accession number</th>
<th>percent coverage</th>
<th>number of peptides</th>
<th>percent share</th>
</tr>
</thead>
<tbody>
<tr>
<td>atToc33</td>
<td>AT1G02280</td>
<td>24</td>
<td>8</td>
<td>1,74</td>
</tr>
<tr>
<td>atTic110</td>
<td>AT1G06950</td>
<td>205</td>
<td>3</td>
<td>0,5</td>
</tr>
<tr>
<td>Actin11</td>
<td>AT3G12110</td>
<td>7,4</td>
<td>3</td>
<td>0,34</td>
</tr>
<tr>
<td>thylakoid lumen 18.3 kDa protein</td>
<td>AT1G54780</td>
<td>37</td>
<td>4</td>
<td>0,9</td>
</tr>
<tr>
<td>Bet v I allergen family protein</td>
<td>AT1G23130</td>
<td>31,5</td>
<td>7</td>
<td>1,34</td>
</tr>
<tr>
<td>LHCA1; chlorophyll binding</td>
<td>AT3G54890</td>
<td>10</td>
<td>4</td>
<td>0,9</td>
</tr>
<tr>
<td>PSBQ; calcium ion binding</td>
<td>AT4G21280</td>
<td>17,5</td>
<td>3</td>
<td>0,7</td>
</tr>
<tr>
<td>(Calmodulin 4); calcium ion binding</td>
<td>AT1G66410</td>
<td>23,5</td>
<td>5</td>
<td>1,2</td>
</tr>
<tr>
<td>calcium-binding protein, putative</td>
<td>AT1G18210</td>
<td>42</td>
<td>6</td>
<td>1,4</td>
</tr>
</tbody>
</table>

### ppi1 NTAPiToc33G

<table>
<thead>
<tr>
<th>protein</th>
<th>accession number</th>
<th>percent coverage</th>
<th>number of peptides</th>
<th>percent share</th>
</tr>
</thead>
<tbody>
<tr>
<td>atToc33</td>
<td>AT1G02280</td>
<td>20</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>atTic110</td>
<td>AT1G06950</td>
<td>2,5</td>
<td>3</td>
<td>0,41</td>
</tr>
<tr>
<td>Actin11</td>
<td>AT3G12110</td>
<td>7,4</td>
<td>3</td>
<td>0,4</td>
</tr>
<tr>
<td>EF1</td>
<td>AT1G07920</td>
<td>7,8</td>
<td>3</td>
<td>0,46</td>
</tr>
<tr>
<td>thylakoid lumen 18.3 kDa protein</td>
<td>AT1G54780</td>
<td>27</td>
<td>6</td>
<td>1,2</td>
</tr>
<tr>
<td>Protein</td>
<td>Accession Number</td>
<td>Percent Coverage</td>
<td>Number of Peptides</td>
<td>Percent Share</td>
</tr>
<tr>
<td>------------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>--------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>atTic110</td>
<td>AT1G06950</td>
<td>8,6</td>
<td>7</td>
<td>1,1</td>
</tr>
<tr>
<td>Actin</td>
<td>AT1G49240</td>
<td>7,4</td>
<td>3</td>
<td>0,3</td>
</tr>
<tr>
<td>EF1</td>
<td>AT1G07920</td>
<td>8</td>
<td>3</td>
<td>0,45</td>
</tr>
<tr>
<td>CPN60A chaperonine</td>
<td>AT2G28000</td>
<td>18</td>
<td>8</td>
<td>1,22</td>
</tr>
<tr>
<td>CPN60B chaperonine</td>
<td>AT1G55490</td>
<td>13,4</td>
<td>6</td>
<td>0,9</td>
</tr>
<tr>
<td>cpHSC70-2</td>
<td>AT5G49910</td>
<td>13</td>
<td>8</td>
<td>1,2</td>
</tr>
<tr>
<td>HSP 93-V</td>
<td>AT5G50920</td>
<td>10</td>
<td>8</td>
<td>1,2</td>
</tr>
</tbody>
</table>
Acknowledgments

First of all I would like to thank my thesis director, Prof. Felix Kessler, for giving me the opportunity to do my PhD thesis in his group at the Laboratory of Plant Physiology in Neuchâtel. I really appreciate your kindness, your motivating enthusiasm and encouragement and thank you for all suggestions and advices.

I would also like to acknowledge Prof. Jean-Marc Neuhaus, Dr. Sacha Baginsky and Dr. Norbert Rolland for accepting to serve in my thesis committee.

A special thank to Dr Birgit Agne for your help, advices and fruitful discussions.

I would like to thank Jana Smutny, Muamba Katambayi, and Sylvain Bischof for technical help.

Un grand merci à tous les membres passés ou présents du laboratoire LPV et en particulier à: Christiane Bettinelli, Claire Bréhélin, Michaël Ilegems, Sibylle Infanger, Christelle Joss, Didier Lochmatter, Meryll Martin, Marlyse Meylan, Joanne Schwaar, Jana Smutny, Pierre-Alexandre Vidi pour leur aide, leur bonne humeur, leur gentillesse et l’excellente atmosphère de travail qui en a découlé.

Je ne remercie pas du tout mais alors vraiment pas les phytotrons et chambres de culture qui ont pris un malin plaisir à tomber en panne très régulièrement, mais je dis un grand merci à M. Jean Pierre Duvoisin pour l’aide apportée dans ces moments là.

Je tiens à remercier également P.J Harvey, Interpol, William Sheller, DominiqueA, ma tablette graphique et les Basler Läckerli qui m’ont accompagné lors de la rédaction de ce manuscrit ; ainsi que Paul-André Siegenthaler et Pia Stieger pour m’avoir fait une place dans leur grand bureau de chefs dès le début de ma rédaction.

Thanx Lalou and Javaid for your support, interest and everything!

Pour finir je souhaite remercier Guillaume : merci pour ton soutien, ton remontage de moral, et pour être là tout simplement (sans oublier de remercier le Dr Gouzerh pour l’aide scientifique !).