

**A molecular genetic study
of Tic20 and Tic22
homologues in *Arabidopsis
thaliana***

Thesis submitted for degree of
Doctor of Philosophy
at the University of Leicester

by
Ali Reza Kasmati, MSc (University of Exeter)
BSc (University of London, Queen Mary College)

Department of Biology
University of Leicester

April 2009

Abstract

In this study *Arabidopsis thaliana* was used as a plant model to investigate the involvement of Tic20 and Tic22 (translocon at the inner envelope membrane of chloroplasts, 20 kD and 22 kD, respectively) in chloroplast protein import. In *Arabidopsis*, there are four Tic20 homologues and two Tic22 homologues, all with predicted similarity to the corresponding pea protein (psTic20 or psTic22). Phylogenetic analyses revealed two, evolutionarily conserved sub-types of both Tic20 and Tic22, termed Group 1 and Group 2 in each case. TargetP analysis was used to predict the subcellular localization of all *Arabidopsis* Tic20 and Tic22 proteins to the chloroplast. Moreover, the TMHMM program was used to identify the transmembrane domains of the atTic20 proteins; all atTic20 homologues have four predicted transmembrane α -helices, like psTic20. To test the TargetP predictions, envelope localization of each protein was tested by transiently expressing YFP fusions in *Arabidopsis* protoplasts. Furthermore, quantitative RT-PCR (and Genevestigator) revealed that all *atTIC20* homologues are expressed throughout development; *atTIC20-I* expression was highest in photosynthetic tissues, whereas *atTIC20-IV* expression was strong in non-photosynthetic tissues and seeds. Quantitative RT-PCR also revealed that *atTIC22-IV* expression is higher than that of *atTIC22-III*. To assess functional significance of the six genes *in vivo*, T-DNA mutants were identified. Homozygous *tic20-I-1* and *tic20-I-2* plants have an albino phenotype correlated with abnormal chloroplast development. To test for functional redundancy, various *tic20* double and triple mutants were studied; apart from those involving *tic20-I*, these were all phenotypically similar to wild type. In contrast, *tic20-I tic20-IV* double homozygotes could not be identified, due to gametophytic and embryonic lethality. Redundancy between atTic20-I and atTic20-IV was confirmed by partial complementation of *tic20-I* by *atTIC20-IV* overexpression. Additionally, *tic22-IV tic22-III* double mutants had a pale phenotype in early plant development, indicating redundancy between atTic22-IV and atTic22-III and a role during early chloroplast biogenesis.

Acknowledgments

I would like to thank my supervisor Prof. Paul Jarvis for his exceptional supervision, patience, encouragement and flexibility throughout this study. I am also grateful to Ramesh Patel, Sybille Kubis, Ghulam Murtaza, Mats Töpel, Jocelyn Bedard, Tony Wardle Sabina Kovacheva, and Morvarid Shir Mohammadi for helping me with their advice and assistance. I also wish to note my appreciation to other members of lab 337, past and present, for their great support (Weihua Huang, Feijie Wu and Qihua Ling). I am also grateful to the Biology Department as a whole for its support. I would like to thank to Natalie Allcock and Stephan Hyman for preparation of samples and images by microscopy. I also thank Reshma Bharkhada from the Department of Genetics for her assistance with the Q-PCR study. I am grateful to the BBSRC for providing me with a research grant to finish this work. I owe special thanks to my advisor, Prof. David Twell, and to my external examiner, Dr. Anil Day, for their expert advice.

Abbreviations

Measurement Units

For the most part, System International (SI) units were employed. Those units used are summarised below.

bp	Base pair (of nucleic acid)
Ci	Curie
Da	Dalton (of protein)
°C	Degree Celsius
g	Gram
<i>g</i>	Unit of force equal to that exerted by gravity
h	Hour
k	Kilo (10 ³)
l	Litre
m	Metre
μ	Micro (10 ⁻⁶)
m	Milli (10 ⁻³)
min	Minute
M	Molar concentration (mol per l)
mol	Mole
n	Nano (10 ⁻⁹)
OD _λ	Optical density (subscript = wavelength in nm)
% (v/v)	Percentage concentration (ml per 100 ml)
% (w/v)	Percentage concentration (g per 100 ml)
rpm	Revolution per minute
s	Second
V	Volt

Chemicals, Solutions and Media

APS	Ammonium persulfate
BSA	Bovine serum albumin
dATP	2'-Deoxyadenosine 5'-triphosphate
dCTP	2'-Deoxycytosine 5'-triphosphate
dGTP	2'-Deoxyguanosine 5'-triphosphate
diH ₂ O	Deionised water
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
dNTP	2'-Deoxyribonucleoside 5'-triphosphate
dTTP	2'-Deoxythymidine 5'-triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetra-acetic acid
EGTA	Ethyleneglycoltetra-acetic acid
EtBr	Ethidium bromide

HEPES acid	<i>N</i> -2-hydroxyethylpiperazine- <i>N'</i> -2-ethanesulfonic acid
IMS	Industrial methylated spirits
IPA	Isopropanol
LB	Luria-Bertani broth
LB-agar	Luria-Bertani agar
MES	2-(<i>N</i> -morpholino)ethanesulfonic acid
MOPS	3-(<i>N</i> -morpholino)propanesulfonic acid
MS	Murashige and Skoog
N _{2(l)}	Liquid nitrogen
SDS	Sodium dodecyl sulphate
TBS	Tris buffer saline
TBS-tween	Tris buffer saline, with tween
TEMED	<i>N,N,N',N'</i> -tetramethylethylenediamine
Tris	Tris (hydroxymethyl) aminomethane

Biological Terms

A	Adenine
aa	Amino acid
CDS	Coding sequence
cDNA	Complementary DNA
cpm	Counts per minute
C	Cytosine
DNA	Deoxyribonucleic acid
EST	Expressed sequence tag
G	Guanine
GFP	Green fluorescent protein
F ₁ , F ₂ , <i>etc.</i>	Filial generation 1, 2, <i>etc.</i>
λ	Lambda
M ₁ , M ₂ , <i>etc.</i>	Mutant generation 1, 2, <i>etc.</i>
MW	Molecular weight
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
RNA	Ribonucleic acid
RT-PCR	Reverse transcriptase PCR
T-DNA	Transfer DNA
T	Thymine
T ₁ , T ₂ , <i>etc.</i>	Transformant generation 1, 2, <i>etc.</i>
U	Uracil
UTR	Untranslated region
u	Unit of enzyme activity
YFP	Yellow fluorescent protein

Table of Contents

A molecular genetic study of Tic20 and Tic22 homologues in <i>Arabidopsis thaliana</i>	1
Abstract	2
Acknowledgments	3
Abbreviations	4

Chapter 1

General Introduction	9
1. The <i>Arabidopsis</i> Model Plant	10
2. Plastids	12
2.1. What are Plastids?	12
2.2. The Origin and Evolution of Plastids	14
2.3. Chloroplast Function and Structure	15
2.4. Chloroplast Protein Import Apparatus	16
2.5. Thylakoid Biogenesis	18
3. Targeting Preproteins to the Chloroplast	19
3.1. Transit Peptides	19
3.2. Cytosolic Factors	20
3.3. Lipid Interactions	22
3.4. Different Stages of Protein Import and Energy Requirements	23
4. The Translocon Complexes	25
4.1. The TOC Complex at the Outer Envelope Membrane (OEM)	25
4.1.1. Toc75	26
4.1.2. Toc159	28
4.1.3. Toc34	30
4.1.4. Toc64	31
4.1.5. Toc12	32
4.2. The TIC Complex at the Inner Envelope Membrane (IEM)	33
4.2.1. Tic110	34
4.2.2. Tic40	35
4.2.3. Tic55	36
4.2.4. Tic32	37
4.2.5. Tic62	38
4.2.6. Tic20, Tic21 and Tic22	38
4.2.7. Stromal Processing Peptidase (SPP)	40
4.2.8. The Chaperones	41
5. Tic20	43
6. Tic22	45
7. Conclusion	47
8. Aims of this project	49

Chapter 2

Materials and Methods	50
2. Molecular Biology	51
2.1. Nucleic Acid and Protein Preparation	51
2.1.1. Plant Genomic DNA Extraction	51
2.1.2. Plant RNA Extraction	52
2.1.3. Crude Plasmid Preparation from Bacterial Overnight Cultures	53
2.1.4. Preparation of High Quality Plasmid DNA	54

2.1.5.	DNA Extraction from Agarose Gels.....	54
2.1.6.	Quantification of DNA and RNA	54
2.2.	Electrophoresis and Related Techniques	55
2.2.1.	Agarose Gels.....	55
2.2.2.	SDS-PAGE	55
2.3.	Bacterial Work.....	57
2.3.1.	Growing Bacteria.....	57
2.3.2.	<i>E. coli</i> (DH5 α) Transformation	57
2.3.3.	<i>Agrobacterium</i> Transformation	58
2.4.	Enzymatic Manipulation of Nucleic Acids.....	59
2.4.1.	Restriction Analysis of Plasmid DNA	59
2.4.2.	Ligation.....	59
2.4.3.	Polymerase Chain Reaction.....	59
2.4.4.	Reverse Transcriptase-PCR.....	60
2.4.5.	Sequencing.....	62
2.5.	<i>Arabidopsis thaliana</i> (ecotype Columbia-0) growth and manipulation	63
2.5.1.	<i>In vitro</i> Culture.....	63
2.5.2.	Soil Culture	63
2.5.3.	Generation of Transgenic <i>Arabidopsis</i>	64
2.5.4.	Cross-Pollination of <i>Arabidopsis</i> Plants	67
2.5.5.	Chlorophyll Measurements.....	67
2.6.	Chloroplast-Related Methods	69
2.6.1.	Chloroplast Isolation.....	69
2.6.2.	Protein Import Reactions	69
2.6.3.	Thermolysin Treatment.....	71
2.7.	Microscopy	72
2.7.1.	Electron Microscopy (EM)	72
2.7.2.	Fluorescence Microscopy	72
2.7.2.1.	Plasmid Constructions	72
2.7.2.2.	Protoplast Preparation.....	76

Chapter 3

Molecular Genetic Study of Tic20 Homologues in <i>Arabidopsis thaliana</i>	77
3.1. Abstract.....	78
3.2. Introduction.....	80
3.3. Results.....	83
3.3.1 Bioinformatics	83
3.3.1.1. <i>Arabidopsis thaliana</i> Tic20 Homologue Protein Alignments	83
3.3.1.2. Comparison with Possibly Related Mitochondrial Proteins and Phylogenetic Analysis.....	85
3.3.1.3. Prediction of Protein Localization (by TargetP).....	90
3.3.1.4. Prediction of Transmembrane Domains (by TMHMM).....	92
3.3.2. Protein Localization by Yellow Fluorescent Protein (YFP).....	95
3.3.3. Genevestigator Analysis	97
3.3.4. Expression Profiles of <i>Arabidopsis</i> TIC20 Homologues.....	99
3.3.5. Mutant Analysis of <i>Arabidopsis</i> TIC20 Genes.....	102
3.3.5.1. Segregation Analysis of <i>Arabidopsis</i> T-DNA Mutants	104
3.3.5.2. Analysis of T-DNA Insertion Mutants by RT-PCR	106
3.3.5.3. Analysis of Albino Phenotype of <i>tic20-I</i> Mutants	108
3.3.5.4. The Analysis of <i>tic20-I</i> Double Mutations	112

3.3.5.5.	Chlorophyll Analysis of Single, Double and Triple Mutants	116
3.3.5.6.	Complementation Analyses	118

Chapter 4

Molecular Genetic Study of Tic22 Homologues in <i>Arabidopsis thaliana</i>	122
4.1. Abstract	123
4.2. Introduction	125
4.3. Results	127
4.3.1. Bioinformatics	127
4.3.1.1. <i>Arabidopsis thaliana</i> Tic22 Homologous Protein Alignments	127
4.3.1.2. Phylogenetic Analysis and Comparison with Other Organisms	130
4.3.1.3. Prediction of Protein Localization (by TargetP)	133
4.3.2. Protein Localization by YFP	134
4.3.3. Protein Import	136
4.3.4. Genevestigator Analysis	138
4.3.5. Expression Profiles of <i>Arabidopsis</i> TIC22 Homologues	140
4.3.6. Mutant Analysis of <i>Arabidopsis</i> TIC22 Genes	142
4.3.5.1. Segregation Analysis of <i>Arabidopsis</i> T-DNA Mutants	144
4.3.5.2. Analysis of T-DNA Insertion Mutants by RT-PCR	146
4.3.5.3. The Analysis of <i>Arabidopsis</i> tic22 Double Mutants	148
4.3.5.4. Chlorophyll Analysis of Single and Double Mutants	148
4.3.7. Summary	153

Chapter 5

General Discussion	154
5.1. Preamble	155
5.2. Tic20	155
5.3. Tic22	158
5.4. Conclusion	162
Appendix 3.1	163
Appendix 3.2	165
Appendix 3.3	170
Appendix 3.4	171
Appendix 4.1	172
References	173

Chapter 1

General Introduction

1. The *Arabidopsis* Model Plant

The *Arabidopsis* plant first was described by Johannes Thal in 1577, and in the last 50 years this plant has been vastly used as a model organism for research purposes. In 2000, the *Arabidopsis* nuclear genomic sequence completed (The-*Arabidopsis*-Genome-Initiative, 2000), which made a great impact and gave many advantages for basic research in genetics and molecular biology. The completed genome and genetic maps of all chromosomes allowed the rapid discovery of many genes in this species, yet, the precise role and function of *Arabidopsis* genes needs to be studied in detail. The high seed yield, ease of growth in small space, rapid development, cross pollination and full fertility of hybrids are some of the advantageous of *Arabidopsis* in academic and industrial research. Furthermore, this plant has a short life cycle that simply could transform to next generation and their progenies are able to be produced in high numbers from a single plant (Somerville and Koornneef, 2002). Over the past decade, numerous laboratories have used a variety of biochemical studies to identify chloroplast protein import machinery components using *Arabidopsis*. The disadvantage of this plant (*Arabidopsis*) compared to other organism such as yeast and mice, is that this plant has extremely low frequency of targeted integration of transforming DNA, whether introduced via *Agrobacterium* or biolistic particles (Puchta, 2002). Despite that, the small size of this plant (*Arabidopsis*) and its parts were regarded as a disadvantage in biochemical study. As T-DNA insertion is nearly random; a few hundred thousand insertions could hit all of the 20,000–30,000 *Arabidopsis* genes more than once. Nevertheless, this plant has become universally recognised as a model plant, which is a non-commercial plant and economically not important as other plants such as tomatoes and potatoes.

There are many resources available throughout the world and here in the UK for *Arabidopsis* research. An enormous number of libraries of insertion mutants could be searched online and seeds could be requested from the *Arabidopsis* Biological Research Centre (ABRC) and Nottingham *Arabidopsis* Stock Centre (NASC). In the UK, the Genomic *Arabidopsis* Resource Network (GARNet) provides access to high throughput genomic technologies, such as metabolic, transcriptomics and proteomics. It is part of the Investigating Gene Function (IGF) initiative funded by the Biotechnology and Biological Sciences Research Council (BBSRC), which aims to characterise gene

functions within *Arabidopsis* and other model organisms with fully sequenced genomes, such as *Drosophila*.

Furthermore, broad ranges of resources are available online which could ease the research for *Arabidopsis* in comparison to other organisms. The National Centre for Biotechnology Information (NCBI) based in Bethesda, Maryland, USA, incorporates Entrez, a text-based search and retrieval system for a variety of different literature, nucleotide and protein sequence, genome, taxonomy and structural databases. The PubMed database is a resource that includes over 14 million citations for biomedical articles dating back to the 1950's. The citations typically have links to full text articles and other related resources, including articles describing *Arabidopsis*. The BLAST (Basic Local Alignment Search Tool) programmes different resource which could be used to find sequences related to a specific polypeptide or nucleotide query (Altschul *et al.*, 1997), and conserved domains can be identified within a protein using their structural search engines. The entire *Arabidopsis* genome sequence can be searched and analysed via this website. The *Arabidopsis* Information Resource (TAIR) is the central point of the online resources for *Arabidopsis*. TAIR provides news and information alongside search engines for insertional mutants, molecular markers, sequence analysis tools, and many external links.

2. Plastids

2.1. What are Plastids?

Plastids are membrane-bound organelles which are found in plant and algae with diverse functions include photosynthesis. This term has originated from the Greek word *plástis* meaning “plastic mouldable”. According to their morphological appearance (colour composition), function and metabolism plastids are generally classified as elaioplasts which are usually found in the endosperm seed, chromoplasts in fruits and petals, amyloplasts in roots, etioplasts in dark-grown seedlings and chloroplasts in photosynthetically active tissues (Neuhaus and Emes, 2000). Figure 1.1 illustrates the structural changes undertaken as the proplastids, or the etioplasts of plants originally grown in the dark, differentiate into mature chloroplasts during leaf development (Whatley, 1978). Thus, plastids have various forms and structure, which switch over depending on the cell type and the environment. As shown in Figure 1.1, the majority of plastids arise from proplastids, which are small organelles found in plants growing in the light as well as dark. Chromoplasts are believed to be formed from chloroplasts, producing carotenoid pigments such as carotene or lycopene (Bouvier *et al.*, 1998). Chromoplasts are pigmented organelles that are found in flowers, fruits, and leaves in stress or senescence. In this stage plastids might senesce not only after chloroplast maturation but also at intermediate stages. The leucoplast is a common name for colourless plastids. Some leucoplasts may contain starch or proteins that act as a storage function (Kirk and Tilney-Bassett, 1978). Like amyloplasts (not always in dark), etioplasts are found in plants growing in the dark (Whatley, 1978). These classifications describe the range of plastid forms and how they could be critical to plant reproduction and development. A more precise classification system for plastids could be based on their physiological and biochemical resources, reflecting the range and forms of plastids which are intermediate between the above classifications. Nonetheless, diverse forms of plastids are present, each with specific structure and unique properties.

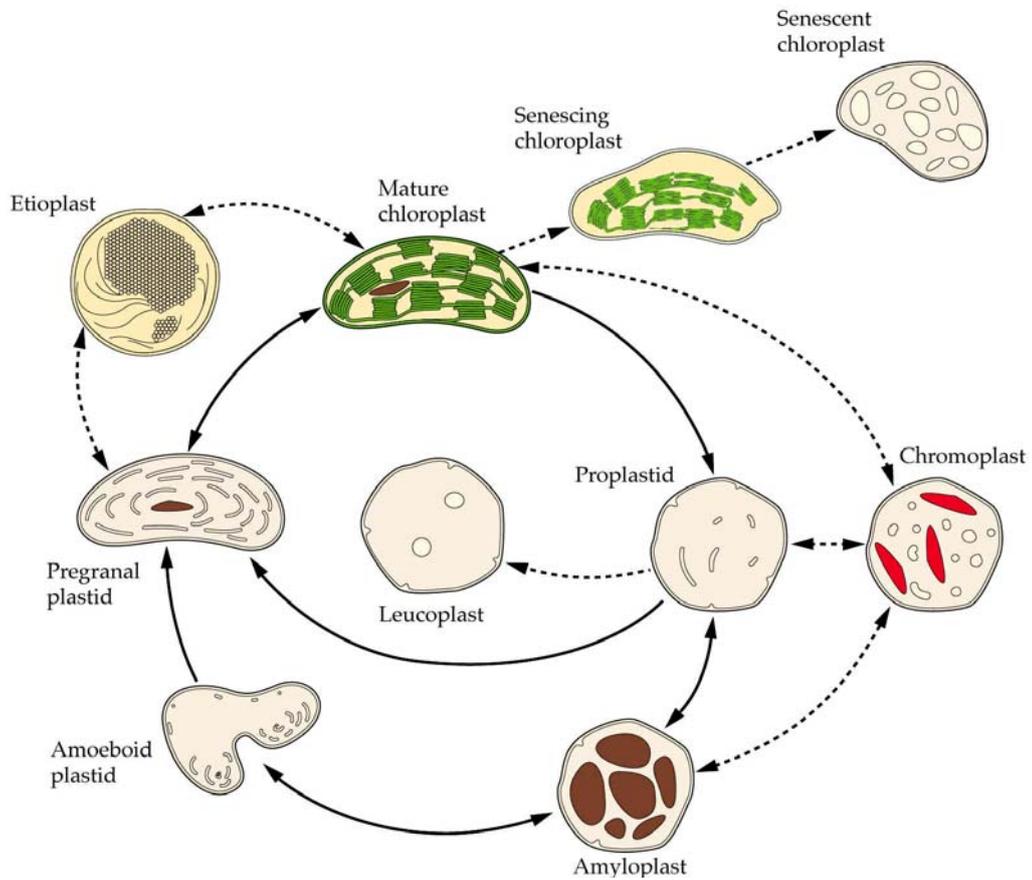


Figure 1.1. The plastid developmental cycles and interrelations between various plastid stages. The solid arrow lines illustrate normal chloroplast development steps; the dotted arrows reveal plastid interrelations that occur under certain environmental or developmental conditions or are unusual (Buchanan *et al.*, 2000) modified from (Whatley, 1978).

2.2. The Origin and Evolution of Plastids

It is generally believed that plastids evolved from a cyanobacterium ancestor, and this is apparent from many lines of evidence including the fact that plastids use part of bacterial division system for their own division thus, information about the mechanism of bacterial division can be used in studies of plastid division (Rothfield and Zhao, 1996; Bramhill, 1997; Errington *et al.*, 2003). Further essential information came from the plastid morphology, biochemistry, genomic organization, and molecular phylogenies of various plastid RNAs and proteins that support their common ancestry (Delwiche *et al.*, 1995; Besendahl *et al.*, 2000). But the exact type of cyanobacterium remains unclear (Mathews *et al.*, 1999). The origin and evolution of plastids is complex and enormous effort has been used to unfold and reveal this complex history. Plastids have their own genome (plastome) that encodes approximately 100 proteins (Hill *et al.*, 1998). (Gray, 1999) and transferred many others to the host nucleus, to ultimately become the specialized organelle it is today (Gray, 1999). Remarkably, the number of protein coding sequences located on the plastome has decreased over evolutionary time (Hill *et al.*, 1998), and evidence indicates that gene transfer may still be taking place. In the *Arabidopsis* nucleus it is estimated that about 4500 genes were transferred from the cyanobacterial endosymbiont during evolution (Martin *et al.*, 2002). Plastids that directly derive from this endosymbiotic event are called “primary plastids”. Such organelles have double envelope membranes that appear to be homologous to the outer and inner membranes of the cyanobacterial endosymbiont (Jarvis and Soll, 2001). However, whether plastid derived from one endosymbiosis or several independent events is disputed.

The majority of algal plastids evolved by “secondary endosymbiosis”, which is an event involving a primary alga being engulfed by another eukaryote, the former eventually falling into the role of an organelle (Archibald and Keeling, 2004). This has give rise to a great diversity of plastids and organisms with plastids. Complex plastids drived from secondary endosymbioses typically have three or more envelope membranes. In general, this process has involved in transferring and losing many genes from the plastids genome, and this has led to the development of new protein targeting systems, which are derived from the secretory system (McFadden, 1999). The great diversity of plastids reveals that plastid evolution is possibly one of the most interesting processes. Because genes can be transferred between distantly related genomes in a

process called lateral or horizontal gene transfer (Keeling, 2004), it is unreliable to rely on all sequence data on their own, and to believe that this reflects the evolution of the organelle. The challenge is to separate those genes that actually reflect the history of the organelle from those that do not.

2.3. Chloroplast Function and Structure

All of plastids are interrelated, which means they have characters which are shared between all of them. Like mitochondria, plastids possess a double membrane with protein complexes that serve as an import apparatus for plastid-targeted proteins. Plastids import and export a wide variety of metabolites and so have selective transport proteins, which are located on the outer and inner envelope membrane (Bendich, 2004). Chloroplasts are plant specific organelles which can be found in all plants' green tissues. They are the most common type of plastids widespread in plant and algae. They play an essential role in photosynthesis and biosynthesis of amino acids and oxygenic (Leister, 2003; Lopez-Juez and Pyke, 2005). Generally, chloroplast converts energy from sunlight into valuable chemical bond that linked redox reactions. Thus, chloroplasts are vital regions for organic and O₂ production which provide fuels necessary for most type of life in this planet (Nelson and Ben-Shem, 2004). This organelle contains the pigment and enzymes necessary for photosynthesis. In contrast to all other plastid variety, mature chloroplasts are the main energy sources. In comparison to mitochondria, chloroplasts are significantly larger and vary in shape, size and number per cell.

Electron microscopy of chloroplasts revealed chloroplasts contain thylakoids stack to form grana (Kirk and Tilney-Bassett, 1978). Each granum consists of a pile of dense vesicles defined one on top other like a pile of coins. These membranes act as a site of electron transport to generate proton gradient in order to produce ATP as an energy source. Starch particles however, are a typical aspect of mature chloroplasts that are an insoluble form of photosynthetic storage. They are surrounded in a ground substance call stroma (or matrix).

Since the *Arabidopsis*, rice and other genomes have been completed, sequence data have played an important role in identification of protein function (Whitelegge, 2003). In addition, the chloroplast genome contributed a significant role in understanding plant genetics and evolution (Golenberg *et al.*, 1993; Clegg *et al.*, 1994;

Morton, 1999). Consequently, chloroplast studies in plant have received significant attention (Schröder and Kieselbach, 2003; Baginsky and Gruissem, 2004; Jarvis, 2004; van Wijk, 2004; Pan *et al.*, 2005). The mature chloroplast is complicated with a large proteome encoded by two different genomes. When a plastid has begun the course towards chloroplast development, numerous unified activities will take place throughout chloroplast biogenesis, all of which are vital for chloroplast functionality.

2.4. Chloroplast Protein Import Apparatus

Initially, protein import was studied using biochemical analyses performed on pea chloroplasts as a model organism, but later switched to *Arabidopsis thaliana* as the model system. The *Arabidopsis* system has numerous advantages compared to other flowering plants, which allowed genomics, genetic analysis and other *in vivo* techniques to be used in chloroplast protein import studies (Jarvis *et al.*, 1998; Bauer *et al.*, 2000; Chen *et al.*, 2002; Chou *et al.*, 2003).

Although, chloroplasts contain fully-functional genetic systems consisting of genomic DNA and the essential transcription and translation machineries, their growth and development depends on proteins encoded in the nuclear genome. The majority of these proteins (>90%) are nuclear encoded and are synthesized in the cytosol as soluble precursors that each have an N-terminal peptide addition called a transit peptide (Schleiff and Klösgen, 2001), and are posttranslationally imported into the plastid (Friedman and Keegstra, 1989). Approximately 2000-3000 proteins are present in mature chloroplasts (Leister, 2003). The transit peptides are essential for chloroplast targeting and translocation of chloroplast across the chloroplast envelope (Keegstra, 1989; De Boer and Weisbeek, 1991; Cline and Henry, 1996). This process is mediated by molecular apparatus in the outer and inner envelope membranes, termed TOC and TIC (Translocon at the outer/inner envelope membrane of chloroplasts), respectively. Upon import, the N-terminal presequence is cleaved at a conserved processing site by the stromal processing peptidase (SPP). The TOC system consists of several protein receptors together with a set of transmembrane proteins that presumably form a protein transport channel across the outer membrane. Previously, several Tic components have been identified: Tic110, Tic62, Tic55, Tic40, Tic32, Tic22, Tic21 and Tic20, but the precise role and function of these proteins in the import process are not well understood

(Küchler *et al.*, 2002). Probably the most crucial function of the TIC complex is the channel formation, but there is a debate about the identity of the inner envelope translocation channel, since it has been suggested that Tic110, Tic20 and Tic21 play roles in channel formation (Kouranov *et al.*, 1998; Heins *et al.*, 2002). There is also evidence that Tic110 serves as a scaffold for stromal chaperones that bind to preproteins as they emerge from the import machinery (Inaba *et al.*, 2003). Although the protein import machinery is well known, the process and the mechanism by which proteins are translocated across either of the two envelope membranes remains in dispute (Lubben *et al.*, 1989).

The main components of the TOC complex are: Toc159, Toc34, and Toc75 (named according to their molecular weights). The Toc complex is involved in the primary recognition of preproteins and their subsequent energy-dependent translocation. Toc75 is a membrane channel protein, which is considered to function as a protein-conducting channel, whereas Toc159 and Toc34 are GTP-binding proteins that function as receptors for transit peptides of chloroplast localized proteins (Kessler *et al.*, 1994; Bauer *et al.*, 2000; Constan *et al.*, 2004; Kubis *et al.*, 2004). The research on pea chloroplasts has been extremely successful, and a large number of TOC and TIC proteins were identified and characterized. Subsequently, based on these studies, several TOC and TIC proteins have been identified in *Arabidopsis* using database search programs, and then studied using different molecular-genetic approaches (Bédard and Jarvis, 2005). The TOC apparatus has also been shown to link with at least three inner membrane proteins. Despite these discoveries, there are many questions relating to the TOC complex which need to be addressed, and there are even more unanswered questions about the Tic complex.

A different group of proteins which perform a vital role in protein import are chaperones. Several chaperone proteins exist which have important functions in chloroplasts. The molecular chaperone Hsp93 is thought to play an important role for the translocation of various proteins across chloroplast membranes, possibly by acting as a motor and consuming energy in the form of ATP hydrolysis. Previously, biochemical and genetic studies suggested that Tic40, Tic110 and Hsp93 cooperate closely during chloroplast protein import (Akita *et al.*, 1997; Nielsen *et al.*, 1997; Stahl *et al.*, 1999; Chou *et al.*, 2003; Inaba *et al.*, 2003; Kovacheva *et al.*, 2005). This protein also assists in other processes, such as protein degradation or disaggregation, when removed from the TIC and present in the stroma (Shanklin *et al.*, 1995; Sokolenko *et*

al., 1998). Chaperones are also involved in providing energy for the posttranslational import of proteins into the endoplasmic reticulum (ER) and mitochondria (Rapaport and Neupert, 1999; Herrmann and Neupert, 2000; Neupert and Brunner, 2002). In contrast with the situation in chloroplasts, heat shock protein 70 (Hsp70) is the major factor which mediates ATP hydrolysis driving import into the ER and mitochondria (Jensen and Johnson, 1999; Pilon and Schekman, 1999; Strub *et al.*, 2000).

2.5. Thylakoid Biogenesis

In both chloroplasts and cyanobacteria the photosynthetic apparatus located on a unique internal membrane system, called thylakoids. Within fully mature chloroplasts, thylakoids are dominating organelle with their unique membrane. The thylakoid membrane structure, formation and composition are closely related to the chloroplasts development from simple undifferentiated proplastids. Chloroplast encoded proteins have essential thylakoid-targeting and assembly signals. As it described earlier, chloroplast require nuclear-encoded proteins to direct translocation across the envelope membranes, and then target proteins to their final destination including into or across thylakoids. Generally, chloroplasts targeting signals fall into two different classes; stroma- and stroma-thylaloid targeting presequences. Although the thylakoid protein import is not fully understood, it has shown that at least three of these pathways have similar ancestry in translocation systems which are analogous to bacteria protein import. These multiple routes discovery provides vital knowledge toward the processes of the photosynthetic protein import machinery. In contrast to cytosolically protein import pathway that use a common protein import apparatus, thylakoid membrane lays on at least four distinguish systems; the Sec, signal recognition particle (SRP), Δ pH/Tat and spontaneous pathways (Bogsch *et al.*, 1998; Chen *et al.*, 2000), which translocation proteins into the thylakoid lumen. Majority of these thylakoidal proteins have huge similarity to protein translocation pathway of the bacterial cell membrane. Like envelope (TOC/TIC of chloroplast) protein import machinery, the thylakoids pathway is essential for the photosynthetic biogenesis. Despite the significance role of thylakoid membrane pathway in photosynthesis and the energy metabolism in plants and cyanobacteria, the molecular processes that associate to the origin, synthesis, control and adaptation of the thylakoids remain unsolved.

3. Targeting Preproteins to the Chloroplast

3.1. Transit Peptides

Proteins localized to the chloroplast contain a short amino-terminal (N-terminal) sequence called a transit peptide (Schleiff and Klösgen, 2001). In eukaryotic cells the “signal hypothesis” has described how secretory proteins are targeted to the endoplasmic reticulum (ER) (Blobel and Sabatini, 1971). Generally, N-terminal targeting signals are used in both prokaryotes and eukaryotes species in order to direct proteins from their site of synthesis to their final destination; these destinations include; the bacterial inner membrane, the ER, mitochondria and chloroplasts in eukaryotic cells (Schatz and Dobberstein, 1996; Schnell and Hebert, 2003). Currently, it is believed that the vast majority of nucleus-encoded proteins of chloroplasts depend upon the presence of an N-terminal transit peptide (Bruce, 2000, 2001). These transit peptides operate in similar style to the cleavable presequences (amino-terminal) of protein import into mitochondria (Bohnert *et al.*, 2007; Neupert and Herrmann, 2007). Although the efficiency of transport could be affected by the mature form of a chloroplast precursor protein (Dabney-Smith *et al.*, 1999; Rial *et al.*, 2002), the transit peptide directly engages the translocation machinery (Sveshnikova *et al.*, 2000b; Hinnah *et al.*, 2002; Inaba *et al.*, 2003; Smith *et al.*, 2004).

Chloroplast transit peptides are generally very different in their length as well as in primary sequence (von Heijne and Nishikawa, 1991). In order to specifically mediate protein import into chloroplast, the chloroplast transit peptides need to share precise primary or secondary structural motifs. In their heterogeneity, transit peptides are incredible (Bruce, 2000, 2001), and there is not enough information available on variations in peptide recognition between the import machinery of higher plant species (Bruce, 2001). Normally, transit peptides share considerable similarity with mitochondrial presequences (Bohnert *et al.*, 2007; Neupert and Herrmann, 2007). Because the two types of targeting sequence are very similar, it is not completely understood how organellar specificity is achieved (Chew and Whelan, 2004; Bhushan *et al.*, 2006). Furthermore, a huge number of preproteins are deliberately “dual-targeted” to both chloroplasts and mitochondria (Silva-Filho, 2003; Duchêne *et al.*, 2005) this revealed that there are significant similarities between the two types of targeting sequence.

Amongst mitochondrial presequences there is no consensus primary structure, whereas in their secondary structure they have a broadly-conserved feature predicted a high proportion of α -helices (Brix *et al.*, 1997; Abe *et al.*, 2000). Interestingly, chloroplast transit peptides do not have secondary structure in aqueous solution (Krimm *et al.*, 1999; Wienk *et al.*, 2000). Moreover, it has been considered that the transit peptide is specially created to form a “perfect random coil” (von Heijne and Nishikawa, 1991), and that this is a vital factor for the recruitment of cytosolic factors which mediate early stages in the import pathway. On the other hand, transit peptides may take on a typical form upon contact with a section of outer envelope membrane; it has been established that certain transit peptides become helical in membrane mimetic environments (Krimm *et al.*, 1999; Wienk *et al.*, 2000; Bruce, 2001). Some studies revealed that the transit peptide interacts effectively with artificial membranes composed of chloroplast lipids *in vitro* (Bruce, 1998). Interestingly, an *Arabidopsis thaliana* mutant deficient in a chloroplast-specific galactolipid (digalacto-syldiacylglycerol (DGDG) displays defects in chloroplast protein import (Chen and Li, 1998). However, the precise role of envelope lipids in the import process remains to be verified such as Monogalactosyldiacylglycerol (MGDG) (Aronsson *et al.*, 2008).

3.2. Cytosolic Factors

Protein import into chloroplasts is a posttranslational process in which soluble cytosolic factors direct precursors from the ribosome to the chloroplast surface (Hiltbrunner *et al.*, 2001; Jackson-Constan *et al.*, 2001). In general, preproteins are threaded through the envelope membrane in unfolded conformation, and so cytosolic chaperones are thought to be essential to avoid their folding or aggregation. Prior study has shown that chloroplast transit peptides are capable of interacting with Hsp70 (Heat shock protein, 70 kDa) chaperones (Jackson-Constan *et al.*, 2001). Previously, proposed that Hsp70, Hsp90 (Heat shock protein, 90 kDa) and 14-3-3 proteins assist the cytosolic steps of chloroplast targeting by forming so-called “guidance complexes” (Qbadou *et al.*, 2006). However, this proposal remains uncertain as this theory is not backed up by *in vivo* studies in *Arabidopsis* (Qbadou *et al.*, 2006).

Chloroplasts have substantial “unfoldase” activity in in that they are capable of unfolding and importing pre-folded precursor proteins (Della-Cioppa *et al.*, 1986;

America *et al.*, 1994); this activity is attributed to chaperone action, pulling on preproteins during import. It is also possible that cytosolic chaperones could be vital for transport of large or hydrophobic preproteins that might target incompetently or aggregate in the cytosol or at the envelope.

Prior study demonstrated (Ivey *et al.*, 2000; Rial *et al.*, 2000; Zhang and Glaser, 2002) that transit peptides are predicted to have high-affinity binding sites for molecular chaperones of the Hsp70 family; these are predicted to mediate direct contacts between these chaperones and transit peptides. In addition, through the import apparatus, including the earliest cytosolic stage of targeting, Hsp70 proteins have been identified as putative protein at different points (Marshall *et al.*, 1990; Schnell *et al.*, 1994; Kourtz and Ko, 1997; May and Soll, 2000). These observations supported the idea that transit peptides are specifically developed to promote chaperone binding efficiency. However, the significance of Hsp70 binding sites in transit peptides remains to be confirmed, as their interruption it seems does not affect protein translocation *in vitro* (Rial *et al.*, 2003, 2006). It is possible that cytosolic factors might do more than sustain preprotein import capacity. A number of transit peptides contain a putative binding site for 14-3-3 proteins which is around a phosphoserine or phosphothreonine residue. It has been shown that the binding of 14-3-3 protein together with Hsp70 and some other components at this site increases the efficiency of chloroplast import *in vitro* (May and Soll, 2000). Since mitochondrial presequences are not capable to form such “guidance complexes”, it has been proposed that this procedure might be beneficial to distinguish transit peptides from structurally similar Mitochondria presequences, so ensuring the fidelity of protein traffic within plant cells (May and Soll, 2000). However, a different study showed that point mutations within the 14-3-3 protein binding site of various preproteins did not have any effect on protein targeting efficiency or specificity *in vivo* (Nakrieko *et al.*, 2004), so the impact of the “guidance complex” hypothesis remains unclear.

This hypothesis however could not be supported by *in vivo* analyses. In both moss (*Physcomitrella patens*) and higher plants (*Arabidopsis*), the knockout mutants lacking the putative receptor, Toc64, were identified, and characterized in extensive detail. In both cases there was no sign of deficiency in plant growth, development or chloroplast protein import (Hofmann and Theg, 2005a; Aronsson *et al.*, 2007). Most importantly, the suggested preprotein of the putative Hsp90/Toc64 transport pathway were shown to be imported with normal efficiency by mutant chloroplasts. Thus, the

importance of the guidance proposal remains to be experimentally proven. It was therefore proposed that Toc64 should be renamed using the more general term, OEP64 (Outer Envelope Protein, 64 kDa) (Hofmann and Theg, 2005c; Aronsson *et al.*, 2007) until such time as its function can be clearly established.

3.3. Lipid Interactions

The envelope membrane are important sites of lipid biogenesis (Douce and Joyard, 1990; Joyard *et al.*, 1998). The glycosyltransferases, MGD1 [MGDG (monogalactosyldia-cylglycerol) synthase 1] and DGD1 [DGDG (digalactosyldiacylglycerol) synthase 1] proteins together form nearly 80% of total chloroplast lipids (Hartel *et al.*, 1997; Bruce, 1998). The chloroplast envelope is the only cytosolically exposed membrane which contains the galactolipids, MGDG and DGDG (Bruce, 1998). Experiments using an *Arabidopsis* DGDG synthase deficient mutant, *dgd1*, have been used to analyse the role of galactolipids in protein import (Chen and Li, 1998). In order for DGDG to be produced, the DGDG synthase needs to catalyse the addition of a second galactose molecule to MGDG. The pale-green phenotype and reduced growth of *dgd1* plants indicates that there is a chloroplast deficiency (Hartel *et al.*, 1997). It has been shown that import in the mutants is defecting at the binding/docking step. Import was shown to be defective at the binding/docking step (that is supported by 100 μ m ATP) (Chen and Li, 1998).

Chloroplast outer envelope proteins are localised with equal efficiency in *dgd1* mutants and wild-type plants. However, preproteins destined for the interior of the plastid are imported less efficiently than in wild-type preproteins (Chen and Li, 1998). Therefore this suggested that DGDG has a role in the early stages of plastid protein import (Chen and Li, 1998). In contrast, the synthesis of MGDG mainly takes place in the chloroplasts inner envelope (Joyard *et al.*, 1983; Miège *et al.*, 1999). Moreover, several *in vitro* studies suggested that MGDG plays an important role during chloroplast protein import (van't Hof *et al.*, 1993; Pinnaduwege and Bruce, 1996; Abe *et al.*, 2000).

The targeting sequences could interact with artificial protein-free membranes (Bruce, 1998), When placed in an environment which mimics the membrane environment they could go through changes that result in major α -helical character. As mentioned before, there is a lot of divergence in primary structure of transit peptides.

thus it is perhaps the secondary structure, induced by the membrane lipids, which facilitates them to be imported into the plastid specifically (Horniak *et al.*, 1993; Wienk *et al.*, 2000). While lipids have a regular composition, transit peptides might have the potential to interrupt and modify the lipid organisation. This also help to allow the preprotein to access the import apparatus (Bruce, 1998).

3.4. Different Stages of Protein Import and Energy Requirements

Generally protein import is thought to involve simultaneous transport through two distinct membranes at “contact sites” (Schnell and Blobel, 1993; Perry and Keegstra, 1994) (Figure 1.2.). The chloroplast protein import can be divided into three different stages based on energy requirements determined *in vitro* with isolated chloroplasts. These steps are thought to reflect the normal series of events which a preprotein follows during import into a plastid. When a preprotein appears on the outer envelope, translocation through the membrane-bound import apparatus will initiate. Initially, preproteins interact with translocon machinery in an “energy-independent binding” step that involves reversible interactions with components of TOC complex; there is no energy consumption at this stage (Perry and Keegstra, 1994; Kouranov and Schnell, 1997). Next, the preprotein becomes deeply inserted into the TOC and also make contact with the TIC machinery (Perry and Keegstra, 1994; Kouranov and Schnell, 1997). This is called an “early import intermediate” and its formation requires low levels of ATP and GTP hydrolysis (Olsen and Keegstra, 1992; Kessler *et al.*, 1994; Young *et al.*, 1999). In the final stage of protein import, the preprotein is entirely translocated into the stroma, where SPP (stromal processing peptidase) is employed to cleave the transit peptide from the precursor. Progression during this step demands high ATP concentrations in the stroma which is consumed by molecular chaperones (Pain and Blobel, 1987; Theg *et al.*, 1989). Similar procedures occur in mitochondria during transport through the outer and inner membrane transport machineries (Horst *et al.*, 1995). In contrast to mitochondrial protein import, chloroplast protein import does not use the proton motive force (Pain and Blobel, 1987; Theg *et al.*, 1989).

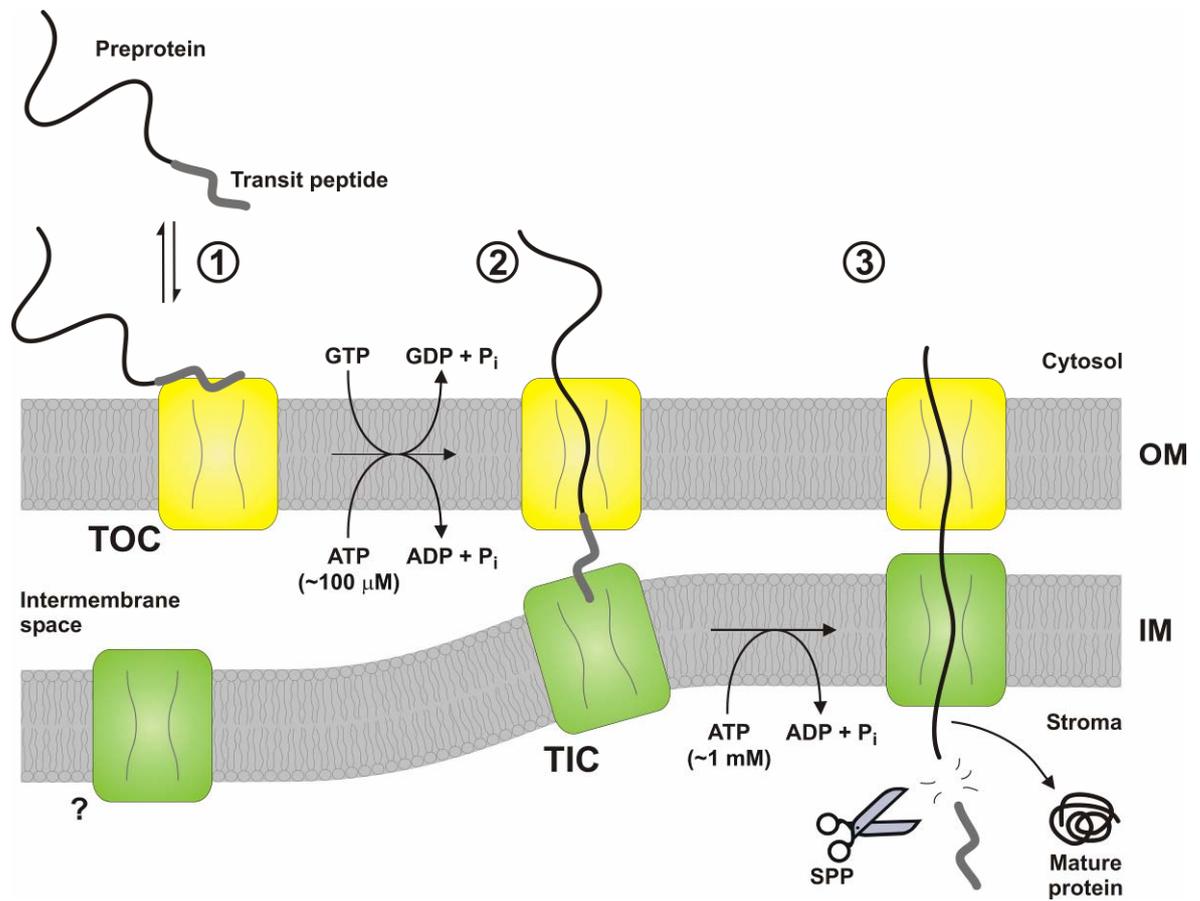


Figure 1.2. Schematic illustration of different stages of the TOC/TIC protein import process. Based on energetic necessities *in vitro*, protein import into chloroplasts can be divided into three different stages. These are related to consecutive steps on the import pathway which happens *in vivo*. At stage 1 (“energy-independent binding”), the preprotein associates reversibly with TOC complex receptor machinery. Stage 2 development (formation of an “early import intermediate”) needs ATP at low concentrations in the intermembrane space, and GTP. At this stage, the preprotein is inserted across the outer envelope membrane (OEM) and is in contact with components of the TIC machinery in the inner envelope membrane (IEM). It is not clear whether the TIC complex exists in a preassembled status, or only forms in response to preprotein arrival in the intermembrane space. Stage 3 (“complete translocation”) needs high ATP concentrations in the stroma. The preprotein is translocated simultaneously across both envelope membranes at a contact site (where the membranes are held in close proximity), the transit peptide is removed by stromal processing peptidase (SPP) (represented by scissors), and the mature protein takes on its final conformation. Figure adapted from Jarvis (2008).

4. The Translocon Complexes

Our understanding of the different translocon complexes which are involved in chloroplast protein import is gradually developing. Initially different laboratories (Hirsch *et al.*, 1994; Kessler *et al.*, 1994; Perry and Keegstra, 1994; Schnell *et al.*, 1994; Wu *et al.*, 1994; Seedorf *et al.*, 1995) used different approaches such as biochemical techniques to characterize the translocon complexes. It has been revealed by two different biochemical methods, cross-linking and co-immunoprecipitation, that different precursor proteins interact directly with proteins of the import machinery, with the TOC and TIC complexes (Schnell *et al.*, 1994). The first method, cross-linking, identified proteins of outer chloroplast membrane called Toc75 and Toc86 during different stages of translocation. Toc86 has been shown to be a fragment of the larger Toc159 protein (Chen *et al.*, 2000). The second method has also recognized Toc34 (thought to be a homologue of a Toc33 in *Arabidopsis*) later it was suggested that the Toc75, Toc159 and Toc34 form the “core” TOC complex (Jelic *et al.*, 2003). According to a standard nomenclature, model all translocon complexes are named based on their location in the outer or inner plastid envelope membrane (Schnell *et al.*, 1997), and the molecular mass of the mature protein, e.g., Toc75 (translocon of the outer chloroplast membrane, 75 kDa). However, the interactions between precursor proteins and chloroplast envelope components during protein import are not well understood, both in their mechanisms and in their sequential order.

4.1. The TOC Complex at the Outer Envelope Membrane (OEM)

The TOC complex machinery consist of Toc34, Toc75 and Toc159, which were first characterized in pea (*Pisum sativum*) using biochemical approaches (Hirsch *et al.*, 1994; Perry and Keegstra, 1994; Schnell *et al.*, 1994; Seedorf *et al.*, 1995; Tranel *et al.*, 1995). The TOC complex is involved in the primary recognition of preproteins and their subsequent energy-dependent translocation. Toc75 is a membrane channel protein, which is considered to function as a protein-conducting channel, whereas Toc159 and Toc34 are GTP-binding proteins that function as receptors for transit peptides of chloroplast localized proteins (Kessler *et al.*, 1994; Bauer *et al.*, 2000; Constan *et al.*, 2004; Kubis *et al.*, 2004). The TOC apparatus has also been shown to link with at least

three inner membrane proteins. It has been reported that the TOC complexes localize to distinct regions of the chloroplast envelope (Subramanian *et al.*, 2001). Recently the entire TOC core complex has been characterized, which is between 500-1000 kDa and consisting of three different proteins (Schleiff *et al.*, 2003b; Kikuchi *et al.*, 2006; Chen and Li, 2007). Stoichiometric analyses have shown that for each Toc159 subunit, there are either three or four Toc34 and Toc75 units (Schleiff *et al.*, 2003b; Kikuchi *et al.*, 2006). Prior study suggested that each TOC complex has a finger-like central site that splits four curved translocation channels (Jelic *et al.*, 2003).

4.1.1. Toc75

Toc75 is a third component of the core complex, and it is deeply embedded within the outer membrane (Perry and Keegstra, 1994). Amino acid sequence analysis has shown that Toc75 contains a β -barrel domain (Schnell *et al.*, 1994; Tranel *et al.*, 1995; Hinnah *et al.*, 1997). In addition, this protein is functionally equivalent to Tom40 protein of mitochondria which also has a β -barrel structure. Toc75 topological studies were based on protease treatments, amino acid sequencing, hydrophobicity, and computer modelling; this revealed that the protein consists of 16 β -strands (Hinnah *et al.*, 1997; Sveshnikova *et al.*, 2000b) or 18 β -strands (Jelic *et al.*, 2003) with an aqueous pore of a small diameter around 14Å–26Å (Hinnah *et al.*, 2002). This indicated that preproteins need to be completely or partially unfolded in order to pass through this pore. It is believed that there are chaperone proteins, such as 14-3-3 and hsp70 homologues, that prevent proteins from folding into their mature shapes, therefore allowing them to pass through the translocon pores (Chen *et al.*, 2000). This pore is comparable in size to that of the Tom40 channel protein of mitochondria, and other proteins translocation pores. Interestingly, prior study suggested that the pore is possibly flexible enough to expand and accommodate small folded proteins (Clark and Theg, 1997). Initially it was considered that the entire length of the Toc75 protein might contribute to channel formation. However, recent studies in bacteria and mitochondria found the Omp85 (Outer membrane protein, 85 kDa) superfamily of proteins, and comparisons led to an alternative hypothesis that the transmembrane β -strands are restricted to a carboxyl-terminal domain. The amino terminus has characteristic POTRA (Polypeptide Transport Associated) domains that are free to perform putative auxiliary functions related to

chaperone activity, transit peptide binding, or TOC complex formation (Sánchez-Pulido *et al.*, 2003; Ertel *et al.*, 2005). Indeed, Toc75 is a founding member of the Omp85 superfamily of components which mediate diverse protein transport processes and are characterized by a similar dual POTRA/ β -barrel domain structure (Reumann *et al.*, 1999; Sánchez-Pulido *et al.*, 2003; Gentle *et al.*, 2005).

The role of Toc75 is not limited to forming the pore through the envelope membrane. It has been shown that pea Toc75 is able to recognize a transit peptide without the support of other TOC components (Hinnah *et al.*, 2002). Interestingly, Toc75 is the only known component of chloroplast outer envelope possessing a transit peptide (Tranel *et al.*, 1995). In fact, it has a bipartite targeting sequence. The transit peptide N-terminus directs the preprotein part way to the stroma, then it is removed by the stromal processing peptidase (Tranel and Keegstra, 1996; Inoue *et al.*, 2001). The C-terminus of the transit peptide possesses a poly-glycine region that is essential for precise targeting to the outer envelope (Davila-Aponte *et al.*, 2003), functioning as an intraorganellar targeting signal.

In *Arabidopsis* there are three closely related homologues of pea Toc75 (psToc75) (Jackson-Constan and Keegstra, 2001). The atToc75-III isoform is the major protein in *Arabidopsis* which has 73% identity to psToc75 in its mature sequence. The other two, atToc75-I and atToc75-IV, were shown to have considerable truncations in their N-termini (Inoue and Potter, 2004). A developmental expression profile of *atTOC75-III* has revealed a similar pattern to that originally reported for psToc75 (Tranel *et al.*, 1995). The mRNA levels in young, developing green tissues are high, and a considerably reduced mRNA levels exist in older tissues (Baldwin *et al.*, 2005). Furthermore, of the two other homologues only *atTOC75-IV* was found to be expressed; *atTOC75-I* found to be a pseudogene due to the presence of a retrotransposon insertion. The *atTOC75-IV* gene is expressed at a particularly low and uniform level during plant growth and in different tissues (Baldwin *et al.*, 2005). Previous results suggest that atToc75-III may be the unique functional orthologue of psToc75 in *Arabidopsis*, and play an important role in chloroplast protein import. Further study supported this, by the finding that a null allele of *atTOC75-III* is embryo-lethal in the homozygous state, suggesting that atToc75-III is crucial for plastid and embryo development (Baldwin *et al.*, 2005). Therefore, Toc75 is believed to be part of a general import apparatus that is common in different plastid-types and all plant species.

4.1.2. Toc159

The Toc159 protein is believed to be the main chloroplast import receptor that interacts with cytosolic precursor proteins at their arrival to the import machinery (Hirsch *et al.*, 1994; Kessler *et al.*, 1994; Perry and Keegstra, 1994) Figure 1.3. There are four Toc159 homologues in *Arabidopsis* (atToc159, atToc132, atToc120, and atToc90) (Hiltbrunner *et al.*, 2001). In *Arabidopsis*, characterization of a *toc159* mutant, termed *ppi2* (plastid protein import), led to the assumption that atToc159 might be a receptor with specificity for in photosynthetic proteins (Bauer *et al.*, 2000). Numerous studies in *Arabidopsis* have tried to address the basis for the existence of multiple Toc-GTPase isoforms. Cross-linking experiments revealed that Toc159 can directly interact with precursor proteins (Perry and Keegstra, 1994; Ma *et al.*, 1996; Kouranov and Schnell, 1997). This interaction is possibly mediated by contact with the N-terminal acidic (A) domain of the Toc159 protein. The large amount of acidic amino acids in the A-domain suggests that this domain interacts with positively charged transit peptides electrostatically. The other two domains of Toc159 are the main GTP-binding (G) domain and the C-terminal membrane anchor (M) domain (Chen *et al.*, 2000). It has considerable sequence identity within its GTP-binding domain to Toc34 (Kessler *et al.*, 1994; Seedorf *et al.*, 1995). The Toc159 and Toc34 receptors are regulated by their GTP-bound state, which control the recognition and translocation cycles (Kessler and Schnell, 2006b). After Toc159 and Toc34 recognition, precursor proteins are then transported to Toc75 which forms the outer envelope translocation channel (Schnell *et al.*, 1994; Tranel *et al.*, 1995; Hinnah *et al.*, 2002). Some data indicate that the Toc159 M domain along with Toc75 forms part of the protein import channel. Therefore, a basic model for Toc159 activity could be the interaction with a precursor protein in the cytosol, followed by Toc159 binding to the TOC complex at the Toc34 protein, and precursor protein import through the membrane channel partially formed by Toc159, and perhaps finally recycling of Toc159 back into the cytosol Figure (1.3).

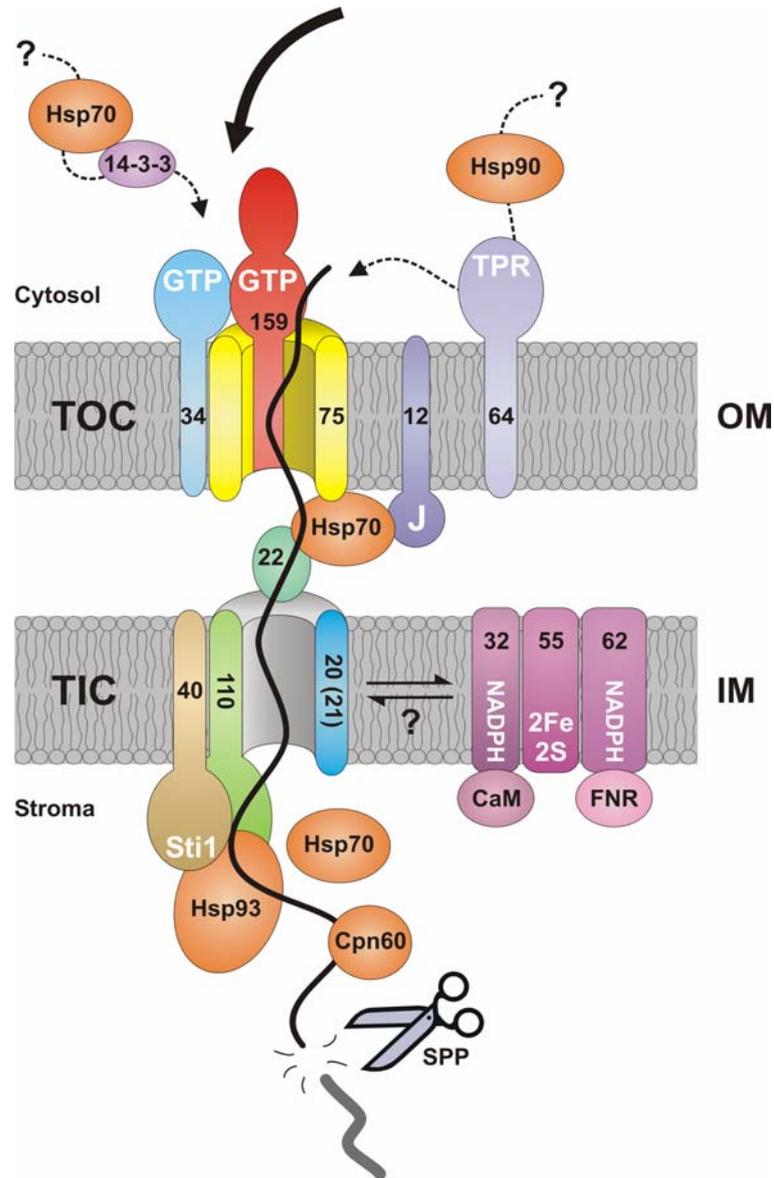


Figure 1.3. The schematic representing the outer envelope membrane and the inner envelope membrane components form the TOC/TIC protein import machinery. Individual components are identified by their predicted molecular weights (represented by number) and some key functional domains are indicated (white text). The core TOC complex is formed by Toc159, Toc34 and Toc75. The first two proteins control preprotein recognition (Toc159, Toc34), and Toc75 forms the outer envelope channel. Cytosolic 14-3-3, Hsp70 and Hsp90 proteins are suggested to form “guidance complexes” that direct preproteins to the TOC apparatus; these may dock at either Toc34 or the peripheral component, Toc64/OEP64, as indicated. Alternatively, protein import may be initiated in the absence of such a guidance complex. It has been suggested that Toc12, Hsp70 and Tic22 act to facilitate the passage of preproteins across the intermembrane space. The inner envelope translocation channel may be formed by Tic110, Tic20 and/or Tic21. The Tic110 protein is also believed to coordinate late events in the process by recruiting stromal molecular chaperones to import sites; in particular, Tic110 has been proposed to collaborate with Tic40 and Hsp93 in a putative stromal import motor complex. Upon arrival in the interior, the transit peptide is cleaved by SPP, and other chaperones (Cpn60 or Hsp70) may assist in the folding or onward transport of the mature domain. Finally, the Tic62, Tic55 and Tic32 components may (in conjunction with ferredoxin NAD(P) reductase (FNR) or calmodulin (CaM)) enable the regulation of import in response to redox signals; these components might only be recruited to import sites under certain conditions, or during the passage of certain preproteins. Figure adapted from Jarvis and Robinson (2004).

4.1.3. Toc34

The TOC translocon complex contains two GTPase proteins, Toc34 and Toc159 (Schleiff *et al.*, 2003a). The G domains of these proteins project into the cytosol recognize and deliver precursor proteins to the translocation pore, Toc75. However, the precise functions of these two GTPase proteins are in dispute (Li *et al.*, 2007). Toc34 is an integral membrane protein that interacts with precursor proteins during protein import and controls preprotein recognition in a GTP-dependent manner (Kessler *et al.*, 1994; Kouranov and Schnell, 1997; Jelic *et al.*, 2002; Sun *et al.*, 2002); it has a short membrane anchor which is a C-terminal transmembrane α -helix (Qbadou *et al.*, 2003). As previously mentioned, the localization of cytosolic Toc159 to the chloroplast is mediated by Toc34 (Wallas *et al.*, 2003). Toc34 like the other major components of the TOC complex was isolated due to its interaction with precursor proteins (Kessler *et al.*, 1994; Schnell *et al.*, 1994; Seedorf *et al.*, 1995). It has also been shown that Toc34 links with Toc75 under non-reducing conditions (Seedorf *et al.*, 1995).

The N-terminal end of the Toc34 protein projects into the cytosol and has typical, conserved motifs of GTP-binding sites, which include the G1 to G4 motif (Bourne *et al.*, 1991; Kessler *et al.*, 1994; Seedorf *et al.*, 1995). Prior study revealed that overexpressed and purified Toc34 protein has endogenous GTPase activity (Seedorf *et al.*, 1995). The GTP hydrolysis rate has been reported to be greatly stimulated by precursor protein (Jelic *et al.*, 2002; Jelic *et al.*, 2003). Toc34, however, is different from other GTPase family members as its G4 domain is not involved in nucleotide specificity or GTPase activity (Aronsson *et al.*, 2003). In contrast to other GTP-binding proteins, Toc34 binds to triphosphate moiety GTP with a higher affinity than the purine base (Jelic *et al.*, 2002). It has been proposed that Toc34 is regulated by phosphorylation (Sveshnikova *et al.*, 2000a; Jelic *et al.*, 2002). In one model for Toc34 action, preprotein recognition involves the precursor binding to unphosphorylated, GTP-bound Toc34, and then hydrolysis of GTP to GDP which releases the precursor towards the translocation channel. Following this release, either Toc34 is phosphorylated to become inactive or it binds to GTP to direct another precursor protein to the translocon channel (Sveshnikova *et al.*, 2000a). However, it is not clear whether Toc34 interacts with preprotein and/or plays a role in the localization of cytosolic Toc159 to the plastid envelope membrane.

4.1.4. Toc64

The pea Toc64 is proposed to act as a receptor; it is an integral protein in the outer membrane, consisting of an inactive amidase or indole acetamide hydrolase domain and three tetratricopeptide repeats (TPR) motif and transmembrane anchor which are typically involved in protein-protein interaction (Sohrt and Soll, 2000). This protein shares some similarity with the Tom70 receptor of mitochondrial protein import receptor. It has been proposed that the Toc64 functions as a docking site for cytosolic guidance complexes consisting of Hsp90 (Qbadou *et al.*, 2006), 14-3-3 and Hsp70 proteins (May and Soll, 2000; Sohrt and Soll, 2000), which transport preproteins. These complexes form in the cytosol and they direct preproteins towards translocation sites. In addition, Toc64 is similar to most chloroplast outer envelope membrane proteins since it does not have a cleavable transit peptide and does not require protease-sensitive components to insert into the membrane.

Under low energy conditions, pea Toc64 could be cross-linked to precursor protein, Toc159, Toc34, and TIC components (Sohrt and Soll, 2000). One of the three essential active site residues normally found in amidase domains, a serine, has been changed to a glycine in pea Toc64 and enzymatic assays detect no amidase activity (Sohrt and Soll, 2000). Additionally, since amidases are soluble proteins, it seems unlikely that Toc64 is an amidase. Based on what is known about Toc64 function, it has been suggested that Toc64 functions early in preprotein translocation, maybe as a docking protein for cytosolic cofactors (Sohrt and Soll, 2000). Not much is known about Toc64 role and function, in contrast to both GTPase proteins.

It has been shown Toc64 double-knockout mutants in *Physcomitrella patens* are indistinguishable from wild type, and have chloroplasts that can import preproteins with normal efficiency (Hofmann and Theg, 2005c). However, it needs to emphasize that Toc64-like proteins in mosses are not necessarily similar to Toc64 in higher plants. Indeed, both Toc64 protein isoforms in moss were found to possess the highly-conserved, catalytic lysine and serine residues of the amidase domain; these are absent in pea Toc64 protein. Therefore, in order to firmly study Toc64 functionality, it was necessary to study mutants in a higher plant species (Aronsson *et al.*, 2007). In *Arabidopsis*, there are three homologous genes (*atTOC64-III*, *atTOC64-V* and *atTOC64-I*) that encode Toc64 related proteins; these proteins were reported to be

localized in chloroplasts, mitochondria and the cytosol, respectively (Chew *et al.*, 2004; Pollmann *et al.*, 2006).

Thus, it has been proposed that Toc64 needs to be renamed using a more common term, OEP64, and that Toc64 function needs to be studied in more detail. Prior study suggested that the Hsp90 chaperone interacts with the TPR domain of Toc64 to deliver preproteins to Toc34 (Schleiff *et al.*, 2003a). However, it has been revealed that the GTP-driven translocation of chloroplast preproteins can occur independently of functional Toc64 (Schleiff *et al.*, 2003a). There are inconsistent reports on Toc64 protein topology (Lee *et al.*, 2004; Hofmann and Theg, 2005c; Qbadou *et al.*, 2007) which has implications for the idea that Toc64 is engaged at both sides of the membrane in protein translocation activities (Becker *et al.*, 2004b; Qbadou *et al.*, 2006).

The topology of Toc64 prediction using various programs revealed this protein have three transmembrane domains (Becker *et al.*, 2004a). These transmembrane regions are essential for correct topology of Toc64 protein (Lee *et al.*, 2004). Prior study revealed that the first region of Toc64 display homology to prokaryotic and eukaryotic amidases, second motif with charge region and third region of C-terminal followed by TPR repeat (Sohrt and Soll, 2000). Recently proposed Toc64 N-terminal transmembrane motif is vital for chloroplasts targeting (Lee *et al.*, 2004). Thus, the exact function of Toc64 in protein translocation is ambiguous (Hofmann and Theg, 2005c; Aronsson *et al.*, 2007).

4.1.5. Toc12

Toc12 is a new component of the TOC complex which exposes a soluble domain into the inner membrane space (IMS). This protein interacts with TOC and TIC components to form supercomplexes (Becker *et al.*, 2004b). The Toc12 protein is a small DnaJ-like co-chaperone which has a C-terminal J-domain which projects into the IMS (Figure 1.3). This leads to the assumption that Toc12 may act as a co-chaperone by controlling the ATPase activity of the IMS Hsp70 (Becker *et al.*, 2004b). In this model, the Toc12 J-domain is associated to Hsp70 protein to stimulate its ATPase activity which in turn promotes the transfer of preproteins to TIC components (Becker *et al.*, 2004b). It is proposed that Toc12 is part of an IMS translocase together with the other components; Hsp70, Toc64 and Tic22 (Figure 1.3) (Becker *et al.*, 2004b).

As discussed earlier, Toc64 is predicted to play a role as a guidance complex receptor. This hypothesis is in dispute, since the role of Toc64 is questionable (Hofmann and Theg, 2005c; Aronsson *et al.*, 2007). At present, there are no *in vivo* studies to back up the role and function of Toc12 (Inoue, 2007). The Toc12 identification revealed a missing gap between the transport machineries (TOC/TIC). However, more data including *in vivo* studies are needed in the future in order to reveal the precise role of this protein in the protein import pathway.

4.2. The TIC Complex at the Inner Envelope Membrane (IEM)

The TIC complexes at the IEM translocate preproteins from the intermembrane space into the stromal compartment. During precursor translocation it has been shown that the TOC and TIC complexes associate (Schnell and Blobel, 1993; Kouranov *et al.*, 1998). The TIC complex also appears to have control of driving precursor proteins into the stroma. The components of the TIC complex have yet to be studied in detail, although several putative components have been identified. It is unclear how many TIC complexes there are, or how the proteins function in protein translocation. Like the TOC components, TIC components were identified in association with precursor proteins in chemical cross-linking and immunoprecipitation experiments (Kessler and Blobel, 1996; Lübeck *et al.*, 1996; Caliebe *et al.*, 1997; Kouranov *et al.*, 1998; Stahl *et al.*, 1999).

The main TIC components identified are: Tic110, Tic62, Tic55, Tic40, Tic32, Tic22, Tic21 and Tic20, but the precise role and function of these proteins in the import process is less well characterized compared to the TOC components (Küchler *et al.*, 2002). Probably the most crucial functions of the TIC complex are channel formation and formation of a motor complex that drive protein import into the stroma; also, there are regulatory factors which control import by recognizing plastid redox status. There is a considerable amount of dispute in the literature over the TIC machinery structure, especially over the main channel component. Since it has been suggested that Tic110, Tic21 and Tic20 play roles in channel formation (Kouranov *et al.*, 1998; Heins *et al.*, 2002; Teng *et al.*, 2006). There is also evidence that Tic110 serves as a scaffold for stromal chaperones that bind to preproteins as they emerge from the import machinery (Inaba *et al.*, 2003).

4.2.1. Tic110

As stated earlier, there is some element of uncertainty on the TIC channel, since there are three different candidates: Tic110, Tic20 and Tic21. Like the TOC components, Tic110 was identified in pea by its close association with preproteins in import at a late stage (Schnell *et al.*, 1994; Wu *et al.*, 1994); this suggests that Tic110 is a major component of TIC complexes (Kessler and Blobel, 1996; Lübeck *et al.*, 1996). Tic110 is an integral IEM protein; its proposed structure consists of two transmembrane helices within its N-terminal membrane anchor domain of ~9 kDa, and a C-terminal domain of ~98 kDa that is mainly hydrophilic. Since Tic110 has a large hydrophilic domain which is resistant to thermolysin, it is believed that this protein is localised mostly in the stroma (Jackson *et al.*, 1998). The stromal location is believed to be vital, since it allows Tic110 to recruit stromal factors needed for import.

Molecular interactions and topology studies of Tic110 have led to two models for the role of this protein: recruiting stromal factors, and a component of the inner envelope protein import channel. Prior study revealed that precursors associate with the soluble region of Tic110 during the late protein import stages; this might act as an anchor for preproteins which could help them to associate with processing factors and chaperones (Inaba *et al.*, 2003). In *Arabidopsis*, the atTic110 protein was shown to be vital for the assembly and function of the chloroplast protein import machinery (Inaba *et al.*, 2005; Kovacheva *et al.*, 2005). Circular dichroism studies revealed that Tic110 forms β -sheets that are typical of pore proteins with a pore diameter of 15 Å and exhibited cation-selectivity (Heins *et al.*, 2002). However, this model was considered dubious by others who suggested a different structure for the C-terminus. Overexpression studies in plants and bacteria revealed that the domain is soluble with an α -helical content; it is proposed to project out into the stromal space (Inaba *et al.*, 2003). Consequently, any Tic110 part that functions as the TIC channel is probably mediated by its N-terminal transmembrane domain, leaving the C-terminus free to control the import process at later stages, a significant component of which seems to be an area that recognizes and binds to TPs (Inaba *et al.*, 2003). Now it is obvious that Tic110 is vital for the import mechanism, as reduced expression causes chlorosis and null mutations are embryo-lethal in *Arabidopsis* (Inaba *et al.*, 2005; Kovacheva *et al.*, 2005).

4.2.2. Tic40

Previously, Tic40 was identified as Toc36 and Cim/Com44 of the chloroplast outer and inner envelope membranes; however, it has since been shown that Tic40 is exclusively an inner membrane protein (Stahl *et al.*, 1999). In pea and *Arabidopsis* this protein is encoded by a single-copy gene (Stahl *et al.*, 1999; Chou *et al.*, 2003). Initially, Tic40 was biochemically characterized in pea, which revealed it to be closely positioned to preproteins at the IEM (Wu *et al.*, 1994; Ko *et al.*, 1995). Cross-linking studies revealed that Tic40 and Tic110 work in close association (Stahl *et al.*, 1999). Tic40 is a co-chaperone that cooperates with Tic110 to recruit and stimulate the ATP hydrolytic activity of ClpC, which is the putative translocation motor for protein import across the TIC channel (Stahl *et al.*, 1999; Chou *et al.*, 2003; Kovacheva *et al.*, 2005; Chou *et al.*, 2006). Similar to other plastid proteins, Tic40 is localized via the general import pathway with the support of a cleavable N-terminal transit peptide.

Tic40 in its N-terminus contains one transmembrane α -helix, with a large C-terminal, hydrophilic domain that protrudes into the stroma (Chou *et al.*, 2003). Tic40 acts as a co-chaperone based on its homology to Hsp70 interacting protein (Hip) and Hsp70-Hsp90 organizing protein (Hop), in its C-terminal domain, termed the Sti1 domain (Chou *et al.*, 2003; Chou *et al.*, 2006). Tic40 also appears to have a TPR protein-protein interaction domain similar to Hip and Hop, which exists upstream of the Sti1 domain (Stahl *et al.*, 1999; Chou *et al.*, 2003). In addition, a recent study demonstrated that the putative Tic40 Sti1 domain could be functionally replaced with the equivalent part of human Hip, which illustrated the significance of a co-chaperone role for Tic40 (Bédard *et al.*, 2007).

In contrast to Tic110, Tic40 appears to be a supporting factor for ClpC function (Stahl *et al.*, 1999; Chou *et al.*, 2003; Kovacheva *et al.*, 2005; Chou *et al.*, 2006), indicating that this component is not essential for protein translocation. This idea is mainly supported by the pale phenotype of the *tic40* T-DNA knockout in *Arabidopsis* (Chou *et al.*, 2003; Kovacheva *et al.*, 2005). For this reason, it is possible that Tic40 function serve to increase the import efficiency.

4.2.3. Tic55

The Tic55 component is believed to be localized at the stromal surface of the chloroplast IEM. It contains a Rieske-type iron-sulphur cluster (normally involved in electron transfer), and a mononuclear iron-binding site that is also located in bacterial oxygenases (Caliebe *et al.*, 1997). Previously, it has been shown that Tic55 is part of TIC complexes in that it interacts with at least five other inner envelope proteins, including Tic110 and ClpC (Caliebe *et al.*, 1997), as well as with Toc86, Toc75 and Toc34. Normally, Tic55 Rieske-type proteins are involved in electron transfer (Caliebe *et al.*, 1997). Based on its cofactors, this subunit is thought to function as a regulatory factor, which acts as biosensor to assemble information in the redox status of the plastid and hence regulate protein import rates through the TIC translocon (Caliebe *et al.*, 1997; Küchler *et al.*, 2002; Hörmann *et al.*, 2004; Chigri *et al.*, 2006; Balsera *et al.*, 2007). Previously, it has been shown in an *in vitro* study in pea diethylpyrocarbonate (DEPC) treatment, an inhibitor that modifies histidine residues, which serve as ligands for Tic55 redox cofactors, can inhibit protein import; therefore it has supported role of Tic55 as part of the TIC complex (Caliebe *et al.*, 1997). However, it is also possible that DEPC was affecting other proteins in the import pathway. In order to verify this hypothesis, *Arabidopsis* Tic55 orthologue (atTic55-II) null mutants were treated with DEPC; equally wild-type and *tic55* mutant chloroplasts were interrupted in terms of import capability (Boij P, Patel R, Jarvis P, Aronsson H, unpublished data).

Thus, Tic55 control of protein import activity according to chloroplast redox signals needs to be tested experimentally in more detail to clarify the role of Tic55. Previously, other groups have failed to determine Tic55 associated within TOC/TIC complexes (Kouranov *et al.*, 1998; Reumann and Keegstra, 1999). Until now, there is not much data available on *Arabidopsis* Tic55 null mutants; Boij *et al.*, have shown that the homozygous null lines of Tic55 are not abnormal compared to the wild type (Boij P, Patel R, Jarvis P, Aronsson H, unpublished data). In *Synechocystis*, a weak homologue of Tic55 was revealed to have a role as a cell death suppressor (Mason and Cammack, 1992).

4.2.4. Tic32

The Tic32 component is localized at the stromal surface of chloroplasts IEM that is essential for chloroplast viability. Tic32 mutant line the putative *Arabidopsis* orthologue of pea TIC32 carrying a T-DNA insertion was studied and proposed Tic32 may be necessary for embryo development. This protein shown to share homology with short-chain dehydrogenase/reductase (SDR), which is analogous to the hydrophobic human retinol dehydrogenases that are bound to the ER membrane (Simon *et al.*, 1999). Tic32 is believed to regulate Ca²⁺/calmodulin dependent activity (Hörmann *et al.*, 2004; Chigri *et al.*, 2006) and might have a dual role in import pathway: first as a regulatory component that possibly affect translocation rates across the IEM (Chigri *et al.*, 2006), second as an crucial subunit that form the entire complex (Hörmann *et al.*, 2004).

After co-immunoprecipitation (Tic22, Tic40, Tic62 and Tic110 for example), it was reported that Tic32 associate with several other Tic components, which believe to play a role in the late translocation process. It was believed calcium regulation occur in IMS or at the IEM, probably engaging TIC components and calmodulin. Previous study of Tic32, employing affinity chromatography confirms that this protein is major IEM protein bound to calmodulin, where interaction was calcium dependent (Chigri *et al.*, 2006). The Tic32, in the same study was revealed to have NADPH-dependent dehydrogenase activity; also NADPH (but not NADH or NADP⁺) influenced the interaction of Tic32 with Tic110. Both NADPH and calmodulin to Tic32 shown to be involved in sensing and combining redox and calcium signals at the TIC complex (Chigri *et al.*, 2006).

Since Tic32 was shown to share homology with SDR activity, and NAD(P) binding site, it was suggested to play a crucial regulatory function. As it was proposed that Tic32 interact with Tic110, it believed that might perform as a redox-regulated gate in TIC channel, in similar way to the β -subunit of potassium channels, with couple oxidoreductase activity to channel inactivation (Bähring *et al.*, 2001).

4.2.5. Tic62

Recently, Tic62 has been characterised as a Tic component interacted with Tic110 and Tic55 (Küchler *et al.*, 2002), also basis on its co-purification with Tic55. The Tic62 N-terminus is localized at the stroma which homology with eukaryotic NAD(H) dehydrogenases and the Ycf390-like proteins in cyanobacteria and non-green algae (Küchler *et al.*, 2002). Circular dichroism studies reveal that Tic62 is build of two structurally different domains (Stengel *et al.*, 2008). With regards to these sequence similarities, Tic62 is believed to function as a protein import regulator by sensing and responding to the organelle redox status (Küchler *et al.*, 2002). As follows, the C-terminus part of Tic62 consist of repetitive module which believe to interact with a ferredoxin-NAD(P)⁺ oxidoreductase (FNR), the photosynthetic pathway enzyme. Through oxygenic photosynthesis, usually FNR mediates electron transfer from ferredoxin to NADP⁺ at the thylakoids site. The Tic62, FNR-binding region of this protein is believe to be a fairly recent development in vascular plants, with no sequence similarity to other known patterns (Balsera *et al.*, 2007); this might be essential for its role in the TIC complex. However, Tic62 localization and its relation with the TIC complex and FNR, are consider to be affect by redox status, including the NADP⁺/NADPH ratio in the stroma (Stengel *et al.*, 2008).

The Tic62 study suggested that these putative substrates could provide valuable within their mode of action in the Tic translocon complex. The Tic62 interaction between Tic110 related or Tic110 unrelated might have transmit information regarding the redox condition of chloroplasts and control protein-import activity in the Tic translocon complex.

4.2.6. Tic20, Tic21 and Tic22

As it mentioned before, the TIC channel characterization is doubtful, since there are three different components; Tic110, Tic20 and Tic21 have been suggested to function on this location (Figure 1.2.). The Tic21 (also called Chloroplast Import Apparatus 5; CIA5), is most recently identified in *Arabidopsis thaliana* as a new member of putative channel component by a forward genetic screen (Teng *et al.*, 2006). The Tic21 protein import study using isolated chloroplast of *tic21* plant, and co-immunoprecipitation

experiments revealed the physical interaction of this protein with the major Toc and Tic complexes (Teng *et al.*, 2006). Therefore, this was indicated that Tic21 is mostly performs at later developmental stages; as it has shown Tic20 express at earlier developmental stage, where Tic21 might take over from Tic20 at later stage. The *cia5* mutant plants exhibit chloroplasts deficiency at an IEM translocation, whereas binding to OEM is unaffected. The null mutant atTic21 plant is albino which accumulates whole preproteins (Teng *et al.*, 2006). Nonetheless, the role of Tic21 has been disputed (Duy *et al.*, 2007), as they supposed that Tic21 functions as an iron carrier and regulator of cellular metal homeostasis. The accumulation of ferritin clusters proof to have different regulation of genes related in iron stress or transport in the mutant. Therefore, it was suggested Permease In Chloroplasts 1 (PIC1) for this protein. Additional analysis required to determine these diverse proposals about Tic21.

Prior study has identified Tic20 and Tic22 as components of protein import machinery, which can be covalently cross-linked to nuclear-encoded preproteins undergoing import across the envelope (Kouranov and Schnell, 1997). They have determined the primary structures of Tic20 and Tic22, and their localization within the inner chloroplast envelope was demonstrated (Kouranov and Schnell, 1997). Tic20 and Tic22 link to other Tic and Toc components to form an active super-complex in the chloroplast envelope membrane (Kouranov *et al.*, 1998). While Tic20 is an integral protein of the inner envelope membrane, Tic22 is peripherally associated to the outer surface of the inner envelope membrane. Tic22 has been suggested to act as a receptor for precursors when they approach from the Toc complex (Kouranov and Schnell, 1997; Kouranov *et al.*, 1999). In *Arabidopsis*, an antisense approach was used to study the importance of *atTIC20-I* for chloroplast biogenesis (Chen *et al.*, 2002). Prior study proposed that Tic20 has four putative transmembrane α -helices (Soll and Schleiff, 2004). An earlier study suggested that Tic20 has three putative transmembrane α -helices (Kouranov *et al.*, 1998). According to my own TMHMM study (First year report), the pea and *Arabidopsis* Tic20 proteins all have four putative transmembrane α -helices. The predicted topology of Tic20 makes it a good candidate for a component of the protein-conducting apparatus of inner membrane translocon. This proposed function is consistent with observed patterns of cross-linking of preproteins to Tic20, which increased at the later stages of protein import when the translocating protein had inserted across both envelope membranes (Kouranov and Schnell, 1997). It has been reported that Tic20 has similarities to bacterial amino-acid transporters and the

mitochondrial Tim17 import component (Rassow *et al.*, 1999). These data suggested that Tic20 functions as a component of the protein-conducting channel at the inner envelope membrane, so far there is no *in vitro* data available to support this idea. In *Arabidopsis*, protein import studies with isolated chloroplasts from control and atTic20-I antisense plants revealed that the antisense chloroplasts are defective mainly in protein translocation across the inner envelope membrane (Chen *et al.*, 2002).

4.2.7. Stromal Processing Peptidase (SPP)

The SPP cleaved off the TP of the preprotein immediately after their arrival in the stroma (Richter and Lamppa, 1999), (Figure 1.2). The SPP is vital in plastid biogenesis, as its down regulation by antisense process triggers albino or lethal phenotypes. Furthermore, SPP reduction levels are influenced largely on the chloroplast import efficiency (Wan *et al.*, 1998; Zhong *et al.*, 2003). This is possibly due to inefficient processing of TOC/TIC apparatus or reflective of a closely incorporated role of SPP, which possibly has an indirect effect at the TIC complex. The SPP N-terminus consist of a zinc-binding motif, which was found in other pitrilysin metalloendopeptidases type, the enzyme responsible for cleaving mitochondrial presequences (VanderVere *et al.*, 1995; Roth, 2004), and it is necessary for catalytic activity (Richter and Lamppa, 2003). The SPP believe to directly interacts with C-terminal ~10-15 residues of the TP, close to the processing site, generally in the area of concentrated basic residues (Richter and Lamppa, 2002). It was believed the SPP composition at cleavage site is rather weak (Emanuelsson *et al.*, 1999), the recent study however, propose that processing involves identification of specific physicochemical properties that is mainly significant than a individual amino acid sequence (Rudhe *et al.*, 2004). The instant release of the mature protein occurs through cleavage of the TP by SPP, where the TP continues to bound to the enzyme for another round of proteolysis, then degraded by a presequence protease (Moberg *et al.*, 2003; Bhushan *et al.*, 2006). The TP could be approached by SPP although the preprotein's C-terminus is connected to the TOC complex, this reveals cleavage happens shortly after entry to the stroma (Schnell and Blobel, 1993). The stromal chaperones; Hsp70 and Cpn60 are understood to mediate the folding of newly imported proteins (Tsugeki and Nishimura, 1993; Jackson-Constan *et al.*, 2001). As Cpn60 was previously interact with the stromal surface of the TIC complex in an ATP-

dependent process, this indicates that folding possibly involve with translocation (Kessler and Blobel, 1996). In contrast, the chaperone possibly binds to the incoming protein mainly to avoid mis-folding or aggregation after completion of translocation, the chaperone-substrate could separate from the TIC to initiate folding in the stroma.

4.2.8. The Chaperones

Chaperones generally regulate the folding or unfolding status of proteins, as well as the assembly or disassembly of multi-subunit complexes. They also supply a driving force for protein translocation across the chloroplast envelope membranes with support of co-chaperones. The 14-3-3 and hsp70 homologues are the chaperone proteins that believe to avoid the proteins from folding to their mature form allowing them to pass via the translocon channels (Chen *et al.*, 2000). Molecular chaperones possibly help protein to progress and reach its final destination, or, the protein might direct to other sub-division of the chloroplast by additional targeting signals. In the literature, it documented the formation of early import intermediates has been called binding and docking. So, in order to preprotein complete translocation, the higher levels of ATP (>100 μM) are needed in the stroma (Theg *et al.*, 1989). This higher energy necessity is recognized to stromal molecular chaperones (Pain and Blobel, 1987). It is understood complete preproteins translocation across OEM and IEM operated by a corresponding chaperone system at the stromal face of the TIC, where substantial energy levels are believe to be required for protein import (Theg *et al.*, 1989). It is mainly believe two stromal molecular chaperones have been relate to the TIC complex: Cpn60 (Kessler and Blobel, 1996), and Hsp93 (Akita *et al.*, 1997; Nielsen *et al.*, 1997) and a plastid homologues of bacterial GroEL (a member of the Hsp60 family) and ClpC (a member of the Hsp100 family), respectively. Formerly Cpn60 was recognized as a major protein by co-immunoprecipitating with Tic110 from solubilized chloroplasts (Kessler and Blobel, 1996), it was later revealed to be associate with TOC-TIC supercomplexes (Kouranov *et al.*, 1998). The lack of other co-immunoprecipitated proteins suggested that Cpn60 may interact directly with the stromal domain of Tic110 (Kessler and Blobel, 1996). These results led to the proposal that Cpn60 is recruited by the Tic110 stromal domain and interact with the incoming unfolded protein that appears from the translocase to support its folding before or as it is released in the stroma.

The second stromal molecular chaperone Hsp93, proven to interact with the TIC complex, was identified by cross-linking experiments in early import intermediates (Akita *et al.*, 1997). As Hsp93 was understood to be a functional homologue of ClpA, the ATPase subunit of the bacterial caseinolytic protease (Clp), it was suggested to have a role in plastid protein degradation (Shanklin *et al.*, 1995). The ClpA in *E. coli* (that is an Hsp100 protein like Hsp93) links with ClpP protease subunit (Schirmer *et al.*, 1996). Both the ClpA and ClpP subunits of the ATP-dependent protease gather into an oligomeric structure composed of two face-to-face ClpP heptameric rings linked to hexameric rings of ClpA at one or both edges (Schirmer *et al.*, 1996). The ClpA protein performs as a protease regulatory subunit particularly by binding and unfolding proteins targeted for degradation in an ATP-dependent manner. Additionally, ClpA could also function separately of the protease subunit, mediating the unfolding protein aggregates, disaggregation. In chloroplasts however, the Hsp93 has been described to be linked to ClpP proteases, providing strong evidence for a role in degradation (Sokolenko *et al.*, 1998; Halperin *et al.*, 2001). Hence, it emerges that in plastids, Hsp93 assists protein translocation, through a TIC complex connection at the IEM, and in protein degradation, via an interaction with ClpP in the stroma. In support of this assumption, translocating preproteins were effectively co-immunoprecipitated from solubilized chloroplasts with Hsp93, and not with other chaperones. This was found that Hsp93 functionally relates to the translocating preproteins, it was also proposed that Hsp100 homologue, rather than an Hsp70 protein in the case of mitochondria, functions as an ATP-dependent protein import motor to drive preprotein translocation.

5. Tic20

The function and formation of the Tic component are not well established compared to Toc complex, however the Tic translocation channel is controversial since at least two candidates have been suggested for this function; the IEM proteins Tic20 and Tic110 (Kouranov *et al.*, 1998; Heins *et al.*, 2002). In cyanobacteria, there are few of Toc/Tic proteins have been identified; Toc75, Tic20, Tic22 and Tic55 (Reumann *et al.*, 1999, 2005). Previously, it shown that psTic20 is an integral membrane protein which has four α -helical transmembrane regions (Chen *et al.*, 2002). In *Arabidopsis*, there are four Homologues: atTic20-I (the isoform is most similar to psTic20), atTic20-IV, atTic20-II and atTic20-V, which are similar to psTic20, where the atTic20-IV, atTic20-II and atTic20-V are less similar to psTic20. However, the psTic20 *in vivo* study revealed that this protein interact with psTic22 (Kouranov and Schnell, 1997).

Phylogenetic studies shown that Tic20 is distantly related to bacterial branched amino acid transporters, including cyanobacterial and mitochondrial putative channel proteins; Tim17 and Tim23 (Reumann *et al.*, 2005). The atTic20 minimal homology and topological similarity with mitochondrial Tim17/22/23 preprotein translocase components support this idea (Reumann *et al.*, 1999, 2005). At the present time there is little known about the chloroplast Tic complex import pathways, the other species might provide some valuable information about this event. In mitochondria there are two Tim complexes; Tim23 and Tim22. The Tim23 is the main part of mitochondria proteins import that forms the channel, which mediates the proteins translocation carrying a N-terminal presequence (Truscott *et al.*, 2001). The Tim22 that is integral inner membrane protein mixing of proteins transport several membrane across hydrophobic helices and internal signal sequences, that metabolite transport of the mitochondrial inner membrane (Kovermann *et al.*, 2002).

Furthermore, the structural correlation and the phylogenetic study of channel-forming proteins convincingly indicated Tic20 as a component of the IEM channel, even though there is lack of evidence *in vitro* and in organellar to back-up this proposal (Soll and Schleiff, 2004). On the basis of the phylogeny analyses, simply atTic20-I and atTic20-IV are likely to be functionally similar to psTic20, albeit a role in protein import cannot be exclude for the remaining atTic20 homologous. In addition the *C. reinhardtii* Tic20 (CrTic20) has shown that this protein is similar to atTic20-I and is therefore likely to be orthologous to psTic20 (Kalanon and McFadden, 2008). The

CrTic20 also contain four predicted transmembrane α -helices (the first helix is weakly predicted), plus a predicted chloroplast targeting transit peptide.

Formerly, Tic22 and Tic20 were identified by chemical crosslinking that shown to be closely related in the intermediate stage with precursor proteins import (Kouranov and Schnell, 1997). Interestingly, using antisense plants the role of Tic20 in protein import was experimentally shown that the level of Tic20 protein was decreased. The antisense plants have pale phenotype, with reduced level of chloroplast proteins and other growth defects at the IEM translocation level, this associated with the Tic20 protein reduced levels (Chen *et al.*, 2002). These data indicates that Tic20 play a significant role for IEM translocation (Chen *et al.*, 2002). Since Tic20 is hydrophobic integral membrane protein, it is however, appropriate to state that this protein might have a role in preprotein conducting channel (Kouranov *et al.*, 1998). In conclusion, the Tic20 antisense plants displayed a strong chlorotic phenotype due to chloroplast development deficiency. In these plants chloroplasts were smaller than wild type plants, surround by interrupt thylakoid membrane developed. It was further verified that these losses were correlated with a reduced level of general chloroplast protein import. In support of Tic20 role as an essential protein in TIC channel, chloroplasts were shown to be particularly incompetent in proteins translocation across the IEM.

6. Tic22

The Tic22 protein was first identified in pea by covalent crosslink experiments as a component of the general protein import apparatus, by its interaction with preproteins trapped at an intermediate stage in the import pathway across the envelope (Kouranov and Schnell, 1997). It has been suggested that Tic22 acts in close correlation with Toc12 and Toc64 (Becker *et al.*, 2004b). In *Arabidopsis*, two homologous genes (*atTIC22-IV* and *atTIC22-III*) encode proteins related to pea Tic22 (psTic22). The psTic22 protein is peripherally associated with the outer surface of the IEM and to some extent interacts with the OEM (Kouranov *et al.*, 1998, Kouranov *et al.*, 1999). It emerged that Tic22 is the first member of TIC complex of intermembrane space that interacts with the outer membrane translocon and therefore direct the preprotein to the inner membrane translocon (Kouranov and Schnell, 1997). Tic22 protein sequence comparison study reveals high sequence similarity of atTic22-IV to psTic22 (60.7% identity), and less similarity of atTic22-III to psTic22 (31.7% identity). Interestingly, Tic22 contains 19% sequence identity over 176 amino acid residues with an open reading frame in the genome of the cyanobacterium *Synechocystis* PCC 6803 (Slr0924) (Reumann and Keegstra, 1999).

The Tic22 protein does not possess any specific sequence motifs (for instance, nucleotide binding domains), neither has it any sequence similarity to other components of protein transport systems. Prior study suggested that Tic22 acts as a receptor for precursors when they approach from the TOC complex (Kouranov and Schnell, 1997). This suggests that this protein might stimulate TOC-TIC supercomplex formation (Figure 1.2). Similar to other components of the TOC and TIC complexes, Tic22 is a nuclear-encoded protein. It is synthesized as a preprotein with 50-amino acid long N-terminal presequence. This presequence directs the protein to its final destination in the IMS.

Deletion mutants and chimeric protein studies revealed that the presequence of Tic22 is required for targeting to the IMS (Kouranov *et al.*, 1999). The protein import of psTic22 was found to be stimulated by ATP and involve of the presence of protease-sensitive components on the chloroplast surface. Import studies using an excess of precursor of the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (pSSU) revealed that Tic22 targeting to the intermembrane space does not engage the general protein import pathway used by stromal preproteins (Kouranov *et al.*, 1999).

This confirmed that the psTic22 presequence does not operate as a stromal transit peptide, and that psTic22 is targeted to the chloroplast intermembrane space by a novel import pathway which is different from known pathways (Kouranov *et al.*, 1999).

7. Conclusion

Comparative studies of chloroplast proteins between *Arabidopsis* and other organisms have identified homologues of which have similar mechanisms to recognise and translocate precursors provide a better understanding about *Arabidopsis* protein import. Important assumptions could be drawn by comparing mitochondrial protein import to chloroplast protein import, as it believed that both organelles have evolved from a bacterial ancestor that originally contained its own genes, which have now been relocated to the nucleus. By comparison of the mitochondrial import pathway, the chloroplast OEM and IEM mainly sit in close association during translocation (Schnell and Blobel, 1993; Schatz and Dobberstein, 1996). Albeit, the mitochondria differ in certain degree, that is require a membrane potential for translocation across the mitochondrial inner membrane (Martin *et al.*, 1991), whereas chloroplast does not (Flügge and Hinz, 1986). Hence, it was concluded that chloroplasts have developed a separate mechanism for directing precursor proteins to the IEM.

So far the core components in the TOC and TIC machinery have been identified. Cross-linking studies of preproteins arrested at different stages of protein import offer valuable information about the roles of different components (Perry and Keegstra, 1994). Research attempts in chloroplast protein import are mainly intended to developing a better understanding of the structural organization of the TOC and TIC complexes. The combination of molecular-genetic studies of *Arabidopsis* and structural development assist to understand chloroplast protein import. Currently, the complete translocations of preprotein procedures from the cytosol to the chloroplast stroma, via the TOC and TIC complexes have been stated in general terms. Nevertheless, the precise aspects of these events remain unidentified. Compared to other protein import event at the TOC complex, the TOC GTPase receptors is better understood than other import aspects. However, the specific method by which the GTP interact between the Toc159 and Toc34 receptors leads to the introduction of preprotein translocation is disputed. The processes, in which these receptor apparatus is able to interact with each other, and the effect of these interactions, need to be determined. Determining how these receptors (Toc34 and Toc159) collaborate to each other to recognize TP signals is necessary to our understanding of how targeting specificity is complete. The TP interactions with different TOC and TIC components present a better understanding of the characteristics of these peptides that are essential for recognition and translocation.

In conclusion, examining the TIC complex and its dynamics expected to provide better understanding about this complex, which support the protein transport across the IEM that remain important. The putative role of Tic20, for example as a component of the import channel requires further study. The multiple isoforms of this protein, present in *Arabidopsis* proposed that like mitochondria, chloroplasts might have different translocon complexes in the IEM which serve to mediate the translocation of different subdivision of preproteins. Additional classification of the putative molecular motors of the IMS and stroma would be required in order to fully recognize how transport across the envelope via the TOC and TIC complexes is motivated.

8. Aims of this project

Chloroplast protein import is an important area of research, because: (a) chloroplasts play various essential roles, including photosynthesis and the biosynthesis of several diverse compounds; (b) there are many unknown details related to the chloroplast protein import apparatus and chloroplast development in general. The overall aim of this study was to expand our knowledge on the roles of two different components of the protein import machinery, Tic20 and Tic22, in particular by making use of the model plant, *Arabidopsis thaliana*. These two TIC complex components are addressed in Chapters 3 and 4, respectively. The specific aims of the work were as follows:

1. To use phylogenetic methods to assess the relationships between the *Arabidopsis* Tic20 or Tic22 homologues and related sequences in other species such as pea and rice.
2. To use fluorescence microscopy to determine the subcellular localization of the *Arabidopsis* Tic20 and Tic22 proteins.
3. To evaluate the developmental and tissue specific expression patterns of the *Arabidopsis* Tic20 and Tic22 genes.
4. To elucidate the *in vivo* roles of the *Arabidopsis* Tic20 and Tic22 genes by identifying and characterizing knockout mutants for each one.
5. To investigate functional relationships between the different Tic20 or Tic22 homologues in *Arabidopsis* by analysing double/triple mutant plants, and by conducting transgenic complementation studies (for Tic20 only).

Chapter 2

Materials and Methods

2. Molecular Biology

All DNA samples were stored at -20°C ; RNA and protein samples were stored at -80°C . Unless otherwise mentioned, all non-refrigerated centrifuge steps were carried out in an Eppendorf centrifuge 5415 D, using a F45-24-11 fixed angle rotor, maximum speed 13,200 rpm ($16,110 \times g$). All refrigerated steps were carried out using a refrigerated Eppendorf centrifuge using a 5417, F-45-30-11 fixed angle rotor, maximum speed 14,000 rpm ($20,817 \times g$).

2.1. Nucleic Acid and Protein Preparation

2.1.1. Plant Genomic DNA Extraction

A rapid DNA extraction method was used to obtain small amounts of *Arabidopsis thaliana* plant genomic DNA suitable for PCR analysis. Approximately 3-4 whole young leaves were collected, placed in a 2.0 ml Safe-Lock microtube (Eppendorf, cat. №: 0030 120.094) with 1-2 sterile glass beads (2.5-3.5 mm diameter) and frozen in N_2 , then stored at -80°C . Previously collected tissue stored at -80°C was ground to a fine powder using the QIAGEN TissueLyser for 1 minute at 30 Hz, without extraction buffer. Once the tissue was ground to a fine powder they were given a 30 s spin (using a refrigerated Eppendorf centrifuge 5417, F-45-30-11 fixed angle rotor, maximum speed 14,000 rpm ($20,817 \times g$), and the powdered samples were suspended in 500 μl extraction buffer (200 mM Tris, pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5% (w/v) SDS) and vortexed briefly for about 10 s to disperse large clumps. Each sample was carried out individually until this point, and then stored on ice until all other samples were ready to proceed to the next stage.

Subsequently, all samples were centrifuged (refrigerated Eppendorf centrifuge 5417, F-45-30-11 fixed angle rotor), at 14,000 rpm ($20,817 \times g$) for 5 min at 4°C and the supernatants were transferred to a fresh 1.5 ml microcentrifuge tube containing 450 μl Isopropyl Alcohol (IPA). The supernatant and IPA was inverted to be mixed and then incubated at -20°C for ~ 1 h or more. After incubation at -20°C , the samples were centrifuged for 10 min at 4°C at 14,000 rpm ($20,817 \times g$). Following centrifugation the supernatant was discarded and 1.0 ml ice cold 70% (v/v) ethanol was added to the pellet

and centrifuged (Eppendorf centrifuge 5417, F-45-30-11 fixed angle rotor) for 5 minutes at 14,000 rpm ($20,817 \times g$). The supernatants were discarded and the pellets were dried for approximately 20-30 min on the bench (at room temperature) and dissolved in 50 μ l sterile diH₂O and kept in fridge (4°C) overnight. After overnight dissolution at 4°C, the samples were centrifuged for 5 min and each supernatant was transferred to a fresh Eppendorf tube. This DNA is stable at -20°C for greater than one year. About 1.0 μ l of this (DNA) sample is sufficient for a standard 20.0 μ l PCR. Using this protocol I have found it possible to process hundreds of individual samples in a single working day. This method is appropriate for *Arabidopsis* plants and has an advantage of not requiring any phenol or chloroform extraction. Thus, it is possible to complete DNA extraction within 15 minutes without handling any hazardous organic solvents (Edwards *et al.*, 1991).

2.1.2. Plant RNA Extraction

Total RNA was extracted from ~100 mg of *Arabidopsis* leaf tissue. The tissue was removed and placed in a sterile 1.5 Eppendorf tube with 1-2 sterile glass beads (2.5-3.5 mm diameter) and placed immediately into N₂(1). Previously collected tissue stored at -80°C was ground to a fine powder using the QIAgen TissueLyser for 1 minute at 30 Hz. All samples were processed individually up to this point and kept on ice. Then the RNeasy[®] Plant Mini Kit (QIAgen, Cat, №: 74904) protocol was used from this point onward according to the manufacturer's instructions.

Purified RNA was eluted in diH₂O and quantified by measuring the wavelengths at 260 nm and 280 nm in a spectrophotometer, using a quartz cuvette. Dilutions of 1.0 in 50.0 were analysed, and RNA concentrations were calculated using the relationship that a 260 nm of 1.0 unit corresponds to 40.0 μ g of RNA per ml of diH₂O. RNA integrity was confirmed by agarose gel electrophoresis.

2.1.3. Crude Plasmid Preparation from Bacterial Overnight Cultures

The alkaline lysis protocol employed was an adaptation of that described by Birnboim and Doly (Birnboim and Doly, 1979), Aliquots (5 ml) of LB (10 g/l yeast extract, 10 g/l peptone, 5 g/l NaCl, pH 7) containing the appropriate antibiotic, in 25 ml universal tubes, were inoculated with single bacterial colonies (*E. coli*) and growth was allowed to proceed overnight at 37°C in a shaking incubator (GallenKamp) at 250 rpm speed. All subsequent steps were carried out in a 1.5 ml Eppendorf tube. Cells (between 1.5 and 3.0 ml overnight culture) were pelleted by centrifugation at 5,000 rpm (2,300 × *g*) for 2 min in a bench-top microfuge (fixed angle rotor Eppendorf centrifuge 5417 D) maximum speed 13,200 rpm (16.100 × *g*).

From this point onward the Qiagen Plasmid Mini Kit (QIAprep[®] Spin Miniprep Kit (Cat. №: 27106) was followed. Then the bacterial pellet was resuspended in 100 µl solution I (50 mM D-glucose, 25 mM Tris-HCl, pH 8.0, 10 mM EDTA). Incubation at room temperature for 5 min was followed by the addition of 200 µl solution II (0.2 M NaOH, 1 % (w/v) SDS); this solution was freshly prepared. All samples were mixed by inversion. Incubation on ice was allowed to proceed for 5 min prior to the addition of 150 µl solution III (3 M potassium, 5 M acetate, pH 4.8; solution III was prepared by adding 11.5 ml glacial acetic acid and 28.5 ml distilled water to 60 ml 5 M potassium acetate). Samples were mixed by inversion, incubated on ice for 5 min, and then centrifuged at 13,200 rpm (16,100 × *g*) for 10 min in a bench-top microfuge (fixed angle rotor; Eppendorf centrifuge 5417 D) maximum speed 13,200 rpm (16.100 × *g*). Supernatants were transferred to fresh tubes containing 1 ml ice-cold ethanol. Samples were mixed by inversion and allowed to incubate on ice for 5 min. Plasmid DNA was collected by centrifugation at 13,200 rpm (16,100 × *g*) for 5 min maximum speed 13,200 rpm (16.100 × *g*), washed with 70 % (v/v) ethanol, dried briefly in a vacuum desiccator, and then resuspended in 40 µl TER (10 mM Tris, 1 mM EDTA, 10 µg/ml RNase).

2.1.4. Preparation of High Quality Plasmid DNA

High quality plasmid DNA was prepared from *E. coli* cultures using the QIAGEN QIAprep Spin Miniprep Kit as well as the QIAGEN Plasmid Midi Kit (QIAprep[®] Spin Miniprep Kit (Cat. №: 27106), following the guidelines provided by the manufacturer.

2.1.5. DNA Extraction from Agarose Gels

DNA was extracted from agarose gels using a QIAex II (Gel Extraction Kit 500 Cat, №: 20051) or QIAquick (Gel Extraction Kit Cat №: 28706) (Qiagen). DNA fragments were resolved by agarose gel electrophoresis as described later (2.2.1). A clean, sharp razor blade was used to excise the required DNA band from the agarose and excess agarose was removed. The DNA was then extracted from the agarose fragment by following the guidelines provided in the QIAex or QIAquick gel extraction kit handbooks provided by the manufacturer. The isolated DNA fragments were then used for cloning or sequencing. For additional information and the buffer recipes refer to the appropriate handbook supplied by the manufacturer.

2.1.6. Quantification of DNA and RNA

Spectrophotometry was used to calculate the approximate concentrations of double stranded DNA (dsDNA) and RNA stock solutions. A sample of the stock solution was diluted either 1:100 or 1:200 in diH₂O. The spectrophotometer was calibrated using diH₂O in a thoroughly rinsed quartz cuvette. Absorbance measurements were made for each sample at wavelengths 260 nm and 280 nm.

Concentrations were calculated using the equations:

$$1 \text{ O.D. unit } 260 \text{ nm} = 50 \text{ ng/}\mu\text{l of dsDNA.}$$

$$1 \text{ O.D. unit } 260 \text{ nm} = 40 \text{ ng/}\mu\text{l of RNA}$$

Alternatively, the concentration of dsDNA samples was estimated by gel quantification as described in 2.2.1.

2.2. Electrophoresis and Related Techniques

2.2.1. Agarose Gels

Agarose gels were employed to resolve DNA and RNA fragments, to determine the concentration of DNA fragments using phage λ DNA standards, and to assess the quality of RNA samples. Loading buffer (10 \times = 0.5% (w/v) Orange G, 50% (v/v) glycerol) was added to the samples to a final concentration of 1 \times prior to loading and electrophoresis. Gels generally contained between 0.8 and 1.5 % (w/v) agarose (depending upon the sizes of fragments to be resolved) and 300 ng/ml EtBr or SYBR (---) safe in 0.5 \times TBE (45 mM Tris, 45 mM Boric Acid, 1 mM EDTA). Agarose was melted in a microwave oven and the solution cooled to \sim 50°C before EtBr or SYBER safe was added. For general analyses, small 10 cm \times 15 cm gels (Wide Mini-Sub Cell; Bio-Rad) were employed. Molecular mass markers were used to determine the size of the DNA fragments. Approximately 0.05-0.1 μ g of 1 kb⁺ DNA ladder (Invitrogen) was added per mm width of well. Gels were run in electrophoresis tanks (Bio Rad) containing 0.5 \times TBE, usually at \sim 10 V/cm. Gels were examined using a short wavelength ultraviolet light transilluminator (UVP) or Syngene and photographed using a using a BioDoc-it™ system (UVP) or Syngene.

2.2.2. SDS-PAGE

Protein separation was performed under denaturing conditions in a discontinuous gel system (Laemmli, 1970). Proteins were separated on a 10%, 12% or 15% acrylamide resolving gel and a 4% stacking gel.

- The following stock solutions were prepared for SDS-PAGE:

- **Resolving (Separating) Gel consisted of:**

	10%	12%	15%
diH ₂ O	3.2 ml	2.67 ml	1.9 ml
1.5 M Tris-HCl (pH 8.8), 0.4% (w/v) SDS	2.0 ml	2.0 ml	2.0 ml
30% (w/v) Acrylamide/0.8% (w/v) Bisacrylamide (Protogel)	2.67 ml	3.2 ml	4.0 ml
10% (w/v) SDS	80.0 µl	80.0 µl	80.0 µl
10% (w/v) APS	40.0 µl	40.0 µl	40.0 µl
TEMED Concentration	8.0 µl	8.0 µl	8.0 µl

- **The Stacking Gel consisted of:**

	4%
diH ₂ O	2.3 ml
0.5 M Tris-HCl (pH 6.8), 0.4% (w/v) SDS	1.0 ml
30% (w/v) Acrylamide/0.8% (w/v) Bisacrylamide (Protogel)	650 µl
10% (w/v) SDS	40.0 µl
10% (w/v) APS	20.0 µl
TEMED	4.0 µl

The stock solutions mentioned above were stable for several months in the refrigerator. For 1.0 l of Coomassie Stain Solution: 1.0 g Coomassie Brilliant Blue 250, R or G 450 ml ethanol, 450 ml H₂O, 100 ml glacial acetic acid. For 1 l Destain Solution: 100 ml methanol, 100 ml glacial acetic acid and 800 ml H₂O.

The samples were loaded onto single gel and electrophoresed in Tris/Glycine/SDS buffer (25 mM Tris, 50 mM glycine, 1% (w/v) SDS) using mini Protean III apparatus (Bio-Rad). Protein gels consist of two phases, the top stacking gel that concentrates the protein to a fine point, and the lower separating gel that separates the proteins according to their size. Prior to gel assembly the glass plates were cleaned with detergent, diH₂O and IMS. First, the separating gel was poured and allowed to set and overlaid with 1 ml diH₂O. After solidification of the separating gel, the overlying solution (diH₂O) was removed and the stacking gel was poured. The stacking phase was poured directly after the addition of TEMED and the combs were inserted. The gel was then allowed to set for at least 15 min at room temperature. Protein samples prepared in

1 × sample buffer (Laemmli, 1970) and heated to 100°C for 2 min before loading into the wells of the gel. Gels were usually run at 150 (V/cm) until the dye front reached the bottom of the gel (~1 h).

Gels were stained for at least 30 min with Coomassie Stain Solution, and then destained for several hours or overnight with regular changes of Destain Solution, until the protein bands were clearly visible, and the surrounding gel was virtually clear. Radioactive gels for autoradiography were dried on filter paper (3M) for 2 h at 80°C using a gel dryer (model № 583 Bio-Rad) attached to a vacuum pump (Vacuubrand, Diaphragm).

2.3. Bacterial Work

2.3.1. Growing Bacteria

Escherichia coli (*E. coli*) (DH5 α) was cultured in LB (LB: 10.0 g of bacto-Tryptone (OXOID, LP0042), 5.0 g bacto-yeast extract, 10.0 g NaCl in 950 ml of diH₂O) in shaking incubator at 37°C (GallenKamp, 250 rpm) or LB-agar (LB: 10.0 g of bacto-Tryptone (OXOID, LP0042), 5.0 g bacto-yeast Extract and 1.5% Bio-Agar added to LB (15.0 g in 1000 ml to solidify the gel), 10.0 g NaCl in 950 ml of diH₂O). When applicable, 100 μ g/ml ampicillin or 50 μ g/ml kanamycin was applied. *Agrobacterium tumefaciens*, strain GV3101::pMP90, was grown in YEP-broth (10 g/l yeast extract, 10 g/l peptone, 5 g/l NaCl, pH 7) in a shaking incubator 30°C (GallenKamp, 250 rpm) or on YEP-agar (10 g/l yeast extract, 10 g/l peptone, 5 g/l NaCl, pH 7, 15.0 g Bio-Agar and 950.0 ml diH₂O) in a bacterial growth cabinet, at 30°C. The following concentrations of antibiotic were applied when applicable; kanamycin 50 μ g/ml, gentamycin 50 μ g/ml, spectomycin 100 μ g/ml.

2.3.2. *E. coli* (DH5 α) Transformation

Aliquots of chemically competent cells (Inoue *et al.*, 1990) in 1.5 ml microfuge tubes were thawed on ice for 10 min. Plasmid DNA (<20 μ l, ~75 ng) was added to 100 μ l of cells, mixed gently, and incubated on ice for 20 min. Cells were subjected to heat shock

at 42°C for 30 s, followed by 2 min incubation on ice. LB (800 µl) was added to each sample prior to incubation at 37°C for 1 h. Cells were pelleted by centrifugation at 10,400 rpm (10,000 × *g*) for 30 s in a bench-top microfuge (fixed angle rotor; Eppendorf centrifuge 5417 D) at maximum speed 13,200 rpm (16.100 × *g*). Most of the supernatant was decanted and the cells were resuspended in the remaining solution (~100 µl) before spreading on LB-Petri-dishes (90 mm) containing the appropriate antibiotic. Plates were incubated upside down overnight at 37°C in order to allow the formation of single colonies. After overnight incubation in the 37°C bacterial growth cabinet, the plates were scored and colonies were picked individually and cultivated.

If pGEM®-T (Promega) plasmid was transformed the plates supplemented with 0.5 mM isopropylthiogalatoside (IPTG) and 80 µg/ml 5-bromo-4-chloro-3-indolyl-β-galactoside (X-Gal). IPTG induces the activity of the β-galactosidase gene that spans the insertion site in the vector. In plasmids with no insert, β-galactosidase is produced, which hydrolyses X-Gal to produce a deep blue indigo compound. If the DNA fragment has successfully inserted into the plasmid, β-galactosidase cannot be produced and therefore no blue colouration is present. After overnight incubation in the 37°C bacterial growth cabinet, the plates were scored and the white colonies picked individually and cultivated.

2.3.3. *Agrobacterium* Transformation

Aliquots of electro-competent cells (50 µl) (McCormac *et al.*, 1998) in 1.5 ml microfuge tubes were thawed on ice for 10 min. Plasmid DNA (~1-5 µg) was added on ice and mixed gently. The cells were incubated on ice for 5 min prior to being flash frozen in N_{2(l)} and thawed at 37°C for 5 min. Then 1.0 ml LB was added to the cells and incubated at 28°C for 2-4 h (GallenKamp, 50 rpm) with gentle shaking. This period allows the bacteria to express the antibiotic resistance genes. Cells were pelleted by centrifugation at 10,400 rpm (10,000 × *g*) for 30 s in a bench-top microfuge (fixed angle rotor; Eppendorf centrifuge 5417 D) at maximum speed 13,200 rpm (16.100 × *g*). Most of the supernatant was decanted and the cells were resuspended in the remaining solution (~100 µl) before spreading on LB plates containing the appropriate antibiotic. The cells were incubated for 2 days at 28°C in a bacterial growth cabinet.

2.4. Enzymatic Manipulation of Nucleic Acids

2.4.1. Restriction Analysis of Plasmid DNA

For general restriction enzyme digest analyses, 0.5-1 µg plasmid DNA was used. Enzymes and buffers were obtained from New England BioLabs, and the recommended buffer was used for each enzyme. At least 3 units restriction endonuclease were used per µg plasmid DNA. The total volume of enzyme used was kept below 10% of the total reaction volume since most restriction enzyme storage buffers contain 50% (v/v) glycerol; reaction buffer glycerol concentrations above 5% (v/v) can result in star activity of the restriction endonuclease (cleavage of sequences similar but not identical to their defined recognition sequences). Digests were generally carried out in a total volume of 20 µl, but volumes were increased when larger amounts of DNA were being digested or restriction endonucleases of low concentration were being used. Digestion was allowed to proceed for at least 60 min (at the temperature recommended by the manufacturer) before analysis using agarose gel electrophoresis.

2.4.2. Ligation

Ligation reactions were performed to introduce DNA fragments into plasmid DNA vector. Ligation reactions contained 1 × buffer supplemented with 1mM ATP, 50 ng of vector, 3 M amount insert DNA and 200-400 units of T4 DNA ligase (New England BioLabs) in a total volume of 10 µl. Ligations were either carried out for 1-4 h at room temperature or overnight at 4°C, and subsequently transformed into *E. coli*. Alternatively, when the pGEM[®]-T Easy Vector System (Promega) was used, procedures were carried out according to the manufacturer's instructions.

2.4.3. Polymerase Chain Reaction

A standard polymerase chain reaction (PCR) was carried out to amplify a specific DNA sequence using oligonucleotide primers, DNA polymerase, nucleotides and a DNA template. The two primers were designed to flank the region of interest and anneal to opposing strands. PCR is a cyclic reaction sequence that involves three steps: template

denaturation, primer annealing to the template, and 5' to 3' extension of the primers by the DNA polymerase (Table 2.1). Each event occurs at a specific temperature. Annealing temperatures were adjusted corresponding to the annealing temperatures of the primers.

Primers were included in reactions at concentrations of approximately 1 μ M each. Each dNTP was included at a concentration of approximately 0.4 mM. *Arabidopsis* genomic DNA templates were added at concentrations of 10-50 ng per 50 μ l reaction; by contrast, 0.01-1 ng of plasmid DNA template was used. The ProofStart DNA Polymerase (2.5 units/ μ l; Qiagen this is a proofreading polymerase which uniquely modified to prevent primer degradation during PCR setup) and Platinum[®] Taq (5 units/ μ l; Invitrogen; this DNA Polymerase is uniquely modified to prevent primer degradation during PCR setup) were used following the guidelines provided by the manufacturers. Amplifications were typically carried out for 30-40 cycles as follows: i) denaturation for 30 s at 94°C; ii) primer annealing for 30 s at temperatures between 55°C and 60°C; and iii) elongation at 68-72°C for 1-3 min depending upon the length of the expected product. Amplifications were generally preceded by an initial template denaturation step at 94°C for 2-5 min, and followed by a final 5 min extension step. Thermal cyclers used were supplied by Biometra (T gradient or T3 thermocycler). When necessary, PCR products were purified by gel electrophoresis and extraction or using a QIAquick PCR purification kit (QIAGEN).

2.4.4. Reverse Transcriptase-PCR

First strand cDNA, complementary to the RNA (5 μ g) previously extracted, was synthesised in a two part reaction. A 12 μ l reaction containing 5 μ g of RNA, 5 μ l of CDS-5' primer (Table 2.1), 1 μ l of 10 mM dNTPs, and diH₂O was incubated at 65°C for 5 min; this led to denaturation of RNA secondary structures. The reaction mixture was then incubated on ice for 2 min to allow annealing the primer. This was followed by the addition of 4 μ l of 5 \times buffer, 2 μ l of 0.1 mM DTT, and 1 μ l of 40 u/ μ l RNase inhibitor (Promega). This 20 μ l reaction was incubated at 42°C in a water bath for an additional 60 min, followed by a reverse transcriptase inactivation step at 70°C for 15 min. During this stage reverse transcriptase elongated the primers to produce the first strand cDNA. PCR was then carried out as described above using the relevant primers. Then, 1 μ l of

cDNA synthesis reaction was used as template in each 25 µl PCR reaction (see Section 2.4.3).

Table 2.1. Primers used for RT-PCR and Q-PCR study.

Gene	Primers Sequences	Primers Melting Temperature
<i>TIC20-I F</i>	5' -CCT CCA GCT TCG ATT ATA CCA TTC C-3'	63.0°C
<i>TIC20-I R</i>	5' -GTT TGT AAC TTT GAC GGC TGC GTT C-3'	63.0°C
<i>TIC20-I R₁</i>	5' -CTT CTT AGT CGT ACG GAA TCT GG-3'	60.6°C
<i>TIC20-I R₂</i>	5' -CAA GCA GCA TAC CCA TCA CTA CAT GG-3'	64.6°C
Q-PCR <i>TIC20-I F</i>	5' -GCA TTT TAG TTG GTA AGG-3'	49.1°C
Q-PCR <i>TIC20-I R</i>	5' -GAG CAA GAT AAT GTT GAG TCT GTC-3'	59.3°C
<i>TIC20-IV F</i>	5' -CTT TCC CAA CGT TGA TTC ATC CGC-3'	62.7°C
<i>TIC20-IV R</i>	5' -GTA TGT AAC ACC CGA AAG GGC CAG-3'	64.4°C
Q-PCR <i>TIC20-IV F</i>	5' -GGT GCG CTC TTG CTG GAG TCT ATG-3'	66.1°C
Q-PCR <i>TIC20-IV R</i>	5' -GTA TGT AAC ACC CGA AAG GGC CAG-3'	64.4°C
<i>TIC20-II F</i>	5' -ATG GCG TCT CTG TGC CTT TCT C-3'	62.1°C
<i>TIC20-II R</i>	5' -CGG TAA TGG TTA ATT AAC AAC ACT G-3'	58.1°C
Q-PCR <i>TIC20-II R</i>	5' -CGG TTG CGC GTC TGC TAG AG-3'	63.5°C
<i>TIC20-V F</i>	5' -TTT GCT CCA TTA CCA TCT CTC AC-3'	58.9°C
<i>TIC20-V R</i>	5' -CAG AAA GCG ATC CAA TAC AGA AG-3'	58.9°C
Q-PCR <i>TIC20-V F</i>	5' -GCT CCA TTA CCA TCT CTC ACC GG-3'	64.0°C
Q-PCR <i>TIC20-V R</i>	5' -CAT CAC CCT TGG ATT GGA GTA CG-3'	62.4°C
<i>TIC22-IV F</i>	5' -GAG TCA TCA GTG AAA CCC AAT C-3'	58.4°C
<i>TIC22-IV R</i>	5' -CCT GCA TGT GTT GTG CAT AAC TTC-3'	61.0°C
Q-PCR <i>TIC22-IV R</i>	5' -GGC GGA GTA GGA AGA GAG AAA C-3'	62.1 °C
<i>TIC22-III F</i>	5' -GTC TCA AGC ATC ATT TTC CCG AG-3'	60.6°C
<i>TIC22-III R</i>	5' -GAG GTT TTA CGA TGC TCC AAG G-3'	60.3°C
Q-PCR <i>TIC22-III R</i>	5' -CGT GTT GGC GTT GAA GAG AGG-3'	59.8°C
Actin-2 F	5' -TCA GAT GCC CAG AAG TCT TGT TCC-3'	62.7°C
Actin-2 R	5' -CCG TAG AGA TCC TTC ATG ATA CC-3'	60.6°C
CDS 5'	5' -TTT TTT TTT TTT TTT TTV N-3'	50.5°C

2.4.5. Sequencing

Sequencing was carried out by the Protein and Nucleic Acid Chemistry Laboratory (PNACL) at the University of Leicester. DNA (1.0 µg) was sent for sequencing in 8.0 µl of diH₂O, along with the appropriate primer (10.0 µl of 1.0 µM). The sequencing results were viewed using DNA Star (Technelysium) Lasergene 7 (Editseq) and Chromas (Technelysium).

2.5. *Arabidopsis thaliana* (ecotype Columbia-0) growth and manipulation

2.5.1. *In vitro* Culture

Arabidopsis seeds for germination *in vitro* were surface-sterilised by sequential immersion and shaking in: (1) 70% (v/v) ethanol, 0.05% (v/v) Triton X-100 for ~5 min; and (2) 100% ethanol for ~10 min. Sterilized seeds were dried on filter papers in a laminar flow hood and sown on MS medium (0.5% (w/v) sucrose, 1× MS salts, 0.05% (w/v) MES, pH 5.7, 0.6% (w/v) agar) (Murashige and Skoog, 1962). When appropriate, the following antibiotics were used: kanamycin (50 µg/ml), phosphinothricin (15 µg/ml), gentamycin (110 µg/ml), hygromycin (15 µg/ml) and sulfadiazine (100 µg/ml). Plates were sealed with micropore tape (3 M) to prevent contamination and allow air exchange. Seed dormancy was broken by stratification in the dark at 4°C from 2 to 4 days. Seedlings were grown at 20-25°C under long day conditions (16 h light, 8 h dark) in Percival growth cabinets (Percival Scientific); light intensity was ~100 µmol photons/m²/s of white light. Plants grown on selective MS medium were generally scored for resistance or sensitivity to the antibiotic after 7-10 days growth. When plants germinated *in vitro* were required to set seed, they were transplanted to soil after 10-14 days and allowed to grow to maturity in the greenhouse.

2.5.2. Soil Culture

Arabidopsis plants were sown in soil composed of Levingtons F2 Seed and Modular Compost, silver sand and vermiculite (medium 2.0-5.0 mm, Sinclair) (proportions ratio 2:0.6:1). Generally, plants were grown at 20°C in a controlled temperature glasshouse under long-day conditions; glasshouse humidity was not regulated. Between October and March, a photoperiod of 16 h was maintained using artificial lighting. Plants were shaded with retractable green netting between May and August. In some cases, particularly when phenotypic observations and growth comparisons were required, plants were grown in growth cabinets (Snijders Scientific Jumo Imago F3000) at 20°C, 16 h photoperiod, and in which the light levels were maintained at ~100 µmol photons/m²/s.

Arabidopsis plants were generally grown in 24 cell compartment trays. Each tray holds 2 l compost, 1 l H₂O, 0.2 l vermiculite and 0.6 l sand. *Arabidopsis* plants germinated *in vitro* were transferred to water-saturated soil after ~2 weeks growth. Plants were kept covered with propagator lids for several days following transferral to soil. *Arabidopsis* seed to be germinated in the glasshouse were sown on soil saturated with water. Trays were covered with an inverted empty tray and placed in a cold room (4°C) for 2-4 days. Following stratification, trays were transferred to the greenhouse. The inverted tray was replaced by a propagator lid to maintain high levels of humidity until seedlings were fully germinated. All plants were watered regularly from the base.

Arabidopsis seed were harvested by rupturing dry siliques over a sheet of paper, and the seed was sieved to remove dry plant material. All newly-harvested seed were allowed to dry at room temperature for several days prior to sowing.

2.5.3. Generation of Transgenic *Arabidopsis*

In this study, cDNAs were amplified by using gene specific primers, and *attB1* and *attB2* primers (Table 2.2), and then the bands were cut from the gel. The bands were purified for cloning purposes and transferred into the pDONR207 vector by the BP reaction, and then transformed into *E. coli* DH5 α strain. Heterozygous *tic20-1-1* plants were transformed by the flora dip method (Clough and Bent, 1998) using an *Agrobacterium*-method. Approximately 20 *Arabidopsis* plants were grown in 10 cm² pots in a growth cabinet or in the greenhouse. More secondary inflorescences could be induced by removing the primary inflorescences when they were ~5 cm in height. Prior to transformation, any young siliques which had formed were removed to increase the chance of transformation. The remaining young secondary inflorescences (no taller than 10 cm with few open flowers) were dipped in solution containing *Agrobacterium* strain GV3101::pMP90.

For each construct, 5 ml LB culture with the appropriate antibiotic was inoculated with a single colony of *Agrobacterium* and then grown overnight to saturation. The next day, 1 ml of overnight culture was used to inoculate 100 ml of fresh LB medium including the appropriate antibiotic. The culture was regularly monitored until it reached an OD₂₆₀ of 1.8 (~14 hours after inoculation). The cells were collected by centrifugation at 5,000 rpm (2038.5 \times g) for 10 min in a GSA rotor (Sorvall[®]

Legend RT) at room temperature. At this point the cells were resuspended in three volumes of infiltration medium (2.2 g/l MS salts, 0.5% (w/v) sucrose, 1 ml/l Gamborg's vitamins, 10 ng/ml 6-BAP (pH adjusted to 5.7). Prior to dipping, 0.05-0.1% (v/v) Silwet (Lehle Seeds) was added and mixed thoroughly. Immediately the culture was poured into a wide plastic pot and *Arabidopsis* inflorescences were submerged in the bacteria solution for 10 min upside down with gentle stirring. Transformation occurs in the ovules of the developing flowers after the female and male gametophyte cell lineages form, but before the embryo develops beyond a single cell (Clough and Bent, 1998). After 10 min, excess liquid was removed on tissue and a plastic sleeve was placed over the plants and folded back to maintain humidity, allowing the bacteria more time in which to enter the buds. Two days after dipping the plastic film was removed and the plants were allowed to set seed. The seed was bulk harvested and the first generation was screened for transformants (T₁). Screening for T₁ seeds was performed on MS medium supplemented with the appropriate antibiotic. In addition, the medium was supplemented with cefotaxime (200 µg/ml), an antimicrobial agent that inhibits the growth of *Agrobacterium*, and benomyl (10 µg/ml), an antifungal agent.

Table 2.2. Primers used for generation of transgenic study.

Gene	Primers Sequences	Primers Melting Temperature
<i>TIC20-IF</i>	5'- <u>AA AAA GCA GGC TCC</u> <u>ATG ATA ACT GGA TAC AGC ACG C</u> -3' attB1 GSP TIC20-I (start)	58.4°C
<i>TIC20-IR</i>	5'- <u>A GAA AGC TGG GTT</u> <u>TTA GTC GTA CGG AAT CTG GAT ATA G</u> -3' attB2 GSP TIC20-I (end)	59.7°C
<i>TIC20-IVF</i>	5'- <u>AA AAA GCA GGC TCC</u> <u>ATG CAG GGT TTG GCG GCG ACC</u> -3' attB1 GSP TIC20-IV (start)	65.7°C
<i>TIC20-IVR</i>	5'- <u>A GAA AGC TGG GTT</u> <u>TCA CCT GAG TGG TCT CTG AAA ACC</u> -3' attB2 GSP TIC20-IV (end)	62.7°C
<i>TIC20-IF</i>	5'- <u>AA AAA GCA GGC TCC</u> <u>ATG GCG TCT CTG TGC CTT TCT CTC</u> -3' attB1 GSP TIC20-II (start)	64.4°C
<i>TIC20-IIR</i>	5'- <u>A GAA AGC TGG GTT</u> <u>TTA GAG TTG TCT ACC GGC GGC ATC</u> -3' attB2 GSP TIC20-II (end)	62.7°C
<i>TIC20-VF</i>	5'- <u>AA AAA GCA GGC TCC</u> <u>ATG GCA ATA ATA TCT CAG</u> -3' attB1 GSP TIC20-V (start)	46.9°C
<i>TIC20-VR</i>	5'- <u>A GAA AGC TGG GTT</u> <u>TCA AAG GAC TTG CCT ATC</u> -3' attB2 GSP TIC20-V (end)	51.4°C

2.5.4. Cross-Pollination of *Arabidopsis* Plants

The *Arabidopsis* plants to be used as parents in genetic crosses were grown *in vitro*, and transferred to soil. Three to four week old plants were used for crossing. Best results were obtained using young, primary inflorescences, although secondary inflorescences were also employed. The flower buds selected for crossing were the most mature of those whose stigmas were still fully enclosed within the sepals. Normally, three to four buds were crossed per female parent, and two to three identical crosses were made (using distinct but genotypically identical plants) for each genetic experiment. Under a dissecting microscope by using ultra-fine forceps, the floral organs, sepals, petals and stamens and anthers were removed leaving the central carpel. The apical meristem and all unwanted buds were carefully removed from the inflorescence. Pollen from donor plants was obtained from young open flowers, then swept against the recipient stigmatic surface in order to transfer pollen. Various flowers were pollinated for each cross, and the inflorescences were carefully labelled. Appearance of an elongated green silique was taken as an indication of a successful cross. Inflorescences carrying F₁ seed were harvested after about three weeks and allowed to dry at room temperature for at least one week prior to sowing. Newly harvested F₁ seed were sown on Petri-dishes (90 mm) with appropriate antibiotic selection. Successful crosses were also confirmed by PCR.

2.5.5. Chlorophyll Measurements

To extract the chlorophyll from plant tissue, samples were placed in DMF in 1.5 ml microcentrifuge tubes (Porra *et al.*, 1989), and mixed overnight on a shaker machine at 4°C. Due to the light sensitivity of the chlorophyll, after collection of the plant tissue and immersion in DMF, the microcentrifuge tubes were masked in a dark environment using aluminium foil. Approximately 15 mg of fresh plant tissue was incubated in 1 ml DMF. Less DMF and more plant tissue was used for plants containing less chlorophyll. After chlorophyll extraction, measurements of the amount of chlorophyll contained within the DMF was made by photometry. The absorbance of the samples was measured at wavelengths 646.8 nm and 663.8 nm. The accuracy of the measurements was maintained by running blanks (DMF) as samples and ensuring that they gave zero values. The concentrations of chlorophyll was calculated according to the equations; Chlorophyll *a* = (13.43 × (A_{663.8} - A₇₅₀)) - (3.47 × (A_{646.8} - A₇₅₀))

$$\text{Chlorophyll } b = (22.9 \times (A_{646.8} - A_{750})) - (5.38 \times (A_{663.8} - A_{750}))$$

Total chlorophyll (nmol) was calculated per mg fresh weight tissue.

2.6. Chloroplast-Related Methods

2.6.1. Chloroplast Isolation

Arabidopsis seeds were sterilized, sown on Petri-dishes (90 mm) and grown for 10 days. Approximately 20-25 Petri-dishes containing ~50-200 seeds each were used for chloroplast isolation. Chloroplasts were isolated according to the procedure described previously (Aronsson and Jarvis, 2002; Kubis *et al.*, 2008). Plant tissue was homogenized by five pulses of ~3-4 s each one in 20 ml of cold chloroplast isolation buffer (CIB) (0.3 M sorbitol, 5 mM magnesium chloride, 5 mM EGTA, 5 mM EDTA, 20 mM HEPES, pH 8.0, 10 mM sodium acetate) using a Polytron® (Kinematica PT 10-35; probe PTA 20S). After each pulse, released chloroplasts were collected in a beaker by filtering the homogenate through a double layer of Miracloth (Calbiochem). The homogenized tissue collected in the Miracloth was dispersed in 20 ml of fresh CIB and homogenized again. The combined homogenate was centrifuged for 5 min at $1000 \times g$ at 4°C and the pellet was resuspended in ~500 µl of CIB. This was loaded onto a linear Percoll gradient and centrifuged in a HB-6 swing-out rotor (Sorvall) at 6910 rpm ($7800 \times g$) at 4°C for 10 min. Intact chloroplasts were recovered and washed once with cold HMS buffer (50 mM HEPES, 3 mM magnesium sulphate, 0.3 M sorbitol). The chloroplast yield was calculated using a haemocytometer (Kubis *et al.*, 2008).

2.6.2. Protein Import Reactions

Template DNA for the *in vitro* transcription/translation of preproteins (atTic22-IV, atTic22-III and preSSU) was amplified by PCR from cDNA clones using M13 primers (Table 2.3) as described previously (Aronsson and Jarvis, 2002). Transcription/translation was performed in a coupled system containing rabbit reticulocyte lysate (TNT T7 Quick for PCR DNA; Promega) or wheat germ extract (TNT Coupled Wheat Germ Extract System; Promega), as well as [³⁵S]methionine and T7 RNA polymerase, according to the manufacturer's instructions (Promega). Identical results were obtained with the different translation systems. Import reactions were carried out in HMS buffer containing 20 mM gluconic acid (potassium salt), 10 mM NaHCO₃ and 20% (w/v) BSA. Each 150 µl import assay contained 10×10^6

chloroplasts, 5 mM MgATP, 10 mM methionine, and translation product not exceeding 10% of the total volume. Import reactions were incubated at 26°C in light for 1-10 min, depending on the experiment. Reactions were stopped with an equal volume of stop buffer (50 mM EDTA, 0.3 M sorbitol, 50 mM HEPES, pH 8.0), and the chloroplasts were pelleted by centrifugation for ~5 s in a bench-top centrifuge (fixed angle rotor; Eppendorf centrifuge 5415 C) at maximum speed 14,000 rpm (16.000 × g).

Table 2.3.

Primers used for the *in vitro* transcription/translation of preproteins.

M13 F	5' -TGT AAA ACG ACG GCC AGT-3'	53.7°C
M13 R	5' -CAG GAA ACA GCT ATG ACC-3'	53.7°C

2.6.3. Thermolysin Treatment

Import reactions were treated with thermolysin to digest any externally localized proteins, since thermolysin cannot penetrate the envelope membranes (Cline *et al.*, 1984). Double sized import reactions (300 μ l) were stopped on ice (stop buffer was not added). The intact chloroplasts were re-isolated by centrifugation at 3,800 rpm (4,200 \times g) for 4 min in a bench-top centrifuge (fixed angle rotor; Eppendorf centrifuge 5415 C). The chloroplasts were very gently resuspended in 300 μ l HMS. At this point the two import reactions were divided equally, one half was treated with thermolysin and the other half was not (control). Thermolysin (100 μ g/ml) and CaCl₂ (300 μ M) were added to the chloroplasts and they were incubated on ice for 30-40 min. Digestion was stopped with an equal volume of EDTA Stop Buffer (0.3 M sorbitol, 50 mM HEPES, 50 mM EDTA, pH 8.0), the samples were pelleted by centrifugation for \sim 5 s in a bench-top centrifuge (fixed angle rotor; Eppendorf centrifuge 5415 C) maximum speed 14,000 rpm (16.000 \times g), and the pellet was resuspended in 2 \times SDS-PAGE sample buffer (See Section 2.2.2).

2.7. Microscopy

2.7.1. Electron Microscopy (EM)

For the studies of leaf and chloroplast structure, cotyledons from plate-grown (10-day-old) seedlings were fixed in 2.5% (w/v) glutaraldehyde, 2% (w/v) paraformaldehyde in 0.1 M sodium cacodylate, pH 6.8, for 2 h at room temperature, and then washed several times in the same buffer. Samples were further fixed using buffered 1% (w/v) osmium tetroxide, washed three times in sterile water, and then dehydrated using an ethanol series and propylene oxide prior to infiltration with Spurr's resin. Once polymerized, both thick (0.5-1.0 μm) and thin (60-80 nm) sections were cut using a Reichert Ultracut S Microtome. The thick sections were collected on glass slides and stained using Toluidine Blue for light microscopy. The thin sections were collected on copper grids, stained with uranyl acetate and lead citrate using a Leica EM Stain grid stainer, and then viewed using a Siemens 102 transmission electron microscope. Electron microscopy analysis was performed with the help of the University of Leicester Electron Microscopy facility.

2.7.2. Fluorescence Microscopy

2.7.2.1. Plasmid Constructions

To study *atTic20* and *atTic22* localization, constructs in which YFP was fused to the 3' end (lacking the stop codon) of the *atTIC20* and *atTIC22* cDNA (Table 2.4) were clones constructed. The YFP-fusion proteins were then analyzed with a fluorescence microscope (Nikon TE-2000E), after introduction of the plasmid constructs by transformation into 10-day-old *Arabidopsis* protoplasts.

The full-length cDNA sequences for the six *Arabidopsis atTIC20* and *atTIC22* genes were identified and ordered from the RIKEN institute (Table 2.4). To generate vectors for fluorescence analysis in transformed plant cells, Gateway entry clones containing the genes were generated in the donor vector, pDONR207, by the BP reaction according to the manufacturer's instructions (Invitrogen). A digestion reaction was employed using *HpaI* and *PstI* restriction enzymes to test that these entry clones carry insertions of the correct size. Also, the entry clones were amplified by PCR using

primers flanking the cloning site (Gent F and Gent R, Table 2.5, C) to further confirm the presence of the insertions. Then, the entry clones were integrated into the YFP fusion vector, p2GWY7 (Karimi *et al.*, 2005), by the LR reaction. The cloning procedure is described in detail below.

In this study, cDNAs were amplified by using gene specific primers, and *attB1* and *attB2* primers (Table 2.5 A and B), and then the bands were cut from the gel. The bands were purified for cloning purposes and transferred into the pDONR207 vector by the BP reaction, and then transformed into *E. coli* DH5 α strain. Transformation of the competent cells occurred with high efficiency resulting in more than 100 colonies, one of which was collected and grown in LB medium overnight at 37°C in a shaking incubator machine [(250 rpm)]. Mini plasmid preparations were made, and then the sizes of the inserts in the entry clones were confirmed by restriction analysis (*HpaI* and *PstI*); digests were separated by agarose gel electrophoresis on 1% agarose gels in 1 \times TBE buffer, stained with ethidium bromide, and then visualized under UV light. The enzyme *HpaI* cleaves at position 307 bp in pDONR207 vector (thus, the distance from the *HpaI* site to the ATG is 125 bp) and *PstI* cleaves at position 2759 bp in pDONR207 vector (thus, the distance from the stop codon to the *PstI* site is 114 bp). Based on this information, the expected fragment sizes were: 1094 bp (*atTIC20-I*), 881 bp (*atTIC20-V*), 1106 bp (*atTIC20-IV*), 1193 bp (*atTIC22-III*) and 1058 bp (*atTIC22-IV*). Digestion of the *atTIC20-II* clone released two fragments of 361 bp and 517 bp because of a second *PstI* site in the Tic20 CDS. Restriction enzyme and sequence analysis of the plasmids indicated that all clones were oriented from 5' to 3' with insertions of the correct size.

Next, using the entry clones, cloning into the YFP destination vector (p2GWY7) (Karimi *et al.*, 2005) was done using the LR reaction (according to manufacturer's instructions; Invitrogen). As before, these constructs were transformed into *E. coli* strain DH5 α to amplify plasmid DNA. One clone was picked up from each transformation for initial evaluation. Plasmid isolations from transformed bacteria were confirmed by PCR as described for the BP reaction.

Table 2.4. Full length cDNAs for the *atTIC20* and *atTIC22* genes.

Gene	Gene Annotation	cDNA clone name	Accession Number	Database	Vector Description	Restriction Enzyme	cDNA Sequences Complete
<i>atTIC20-I</i>	At1g04940	RAFL16-70-E19	AK117165	RIKEN	pBluescript II SK*	XhoI/BamHI	Confirmed 2007
<i>atTIC20-IV</i>	At4g03320	RAFL05-07-O03	AY040017	RIKEN	pBluescript II SK*	XhoI/SacI	Confirmed 2007
<i>atTIC20-II</i>	At2g47840	RAFL05-17-E19	AY050346	RIKEN	pBluescript II SK*	XhoI/SacI	Confirmed 2007
<i>atTIC20-V</i>	At5g55710	RAFL05-13-F24	BT001996	RIKEN	pBluescript II SK*	XhoI/SacI	Confirmed 2007
<i>atTIC22-IV</i>	At4g33350	RAFL21-14-F06	AK118805	RIKEN	pBluescript II SK*	XhoI/BamHI	Confirmed 2007
<i>atTIC22-III</i>	At3g23710	151A14	T76759	ABRC	pZL1	Sall/NotI	Confirmed 2007

* Modified

Table 2.5. (a, b and c) Primers used for C-Terminal YFP fusion study.

(a)

Gene	Primers Sequences	Primers Melting Temperature
<i>TIC20-I</i> F	5' - <u>AA AAA GCA GGC TCC</u> <u>ATG ATA ACT GGA TAC AGC ACG C</u> -3' attB1 GSP <i>TIC20-I</i> (start)	58.4°C
<i>TIC20-I</i> R	5' <u>A GAA AGC TGG GTT</u> <u>GTC GTA CGG AAT CTG GAT ATA GG</u> -3' attB2 GSP <i>Tic20-2</i> (end)	60.6°C
<i>TIC20-IV</i> F	5' - <u>AA AAA GCA GGC TCC</u> <u>ATG CAG GGT TTG GCG GCG ACC</u> -3' attB1 GSP <i>TIC20-IV</i> (start)	65.7°C
<i>TIC20-IV</i> R	5' - <u>A GAA AGC TGG GTT</u> <u>CCT GAG TGG TCT CTG AAA ACC TCC</u> 3' attB2 GSP <i>TIC20-IV</i> (end)	64.4°C
<i>TIC20-II</i> F	5' - <u>AA AAA GCA GGC TCC</u> <u>ATG GCG TCT CTG TGC CTT TCT CTC</u> -3' attB1 GSP <i>TIC20-II</i> (start)	64.4°C
<i>TIC20-II</i> R	5' - <u>A GAA AGC TGG GTT</u> <u>GAG TTG TCT ACC GGC GGC ATC AG</u> -3' attB2 GSP <i>TIC20-II</i> (end)	66.0°C
<i>TIC20-V</i> F	5' - <u>AA AAA GCA GGC TCC</u> <u>ATG GCA ATA ATA TCT CAG</u> -3' attB1 GSP <i>TIC20-V</i> (start)	46.9°C
<i>TIC20-V</i> R	5' - <u>A GAA AGC TGG GTT</u> <u>AAG GAC TTG CCT ATC AGC</u> -3' attB2 GSP <i>TIC20-V</i> (end)	53.7°C

(b)

<i>TIC22-IV</i> F	5' - <u>AA AAA GCA GGC TCC</u> <u>ATG GAG TCA TCA GTG AAA C</u> -3' attB1 GSP <i>TIC22-IV</i> (start)	52.4°C
<i>TIC22-IV</i> R	5' - <u>A GAA AGC TGG GTT</u> <u>CTC TTT GAT CAA ATC CTG</u> -3' attB2 GSP <i>TIC22-IV</i> (end)	49.1°C
<i>TIC22-III</i> F	5' - <u>AA AAA GCA GGC TCC</u> <u>ATG AAT TCA AAC ATT TTC CCA CC</u> -3' attB1 GSP <i>TIC22-III</i> (start)	55.3°C
<i>TIC22-III</i> R	5' - <u>A GAA AGC TGG GTT</u> <u>CTC CTG TGT TTG CTC AGT TG</u> -3' attB2 GSP <i>TIC22-III</i> (end)	57.3°C

(c)

Gent F	5' -CAT TTC ATT TGG AGA GGA CTC CG-3'	60.6°C
Gent R	5' -CGG ACA CGC TGA ACT TGT GGC-3'	63.7°C

2.7.2.2. Protoplast Preparation

The above constructions were introduced into wild-type *Arabidopsis* protoplasts (Columbia-0 ecotype), prepared as follows. Green leaves of 10-day-old *Arabidopsis* seedlings grown *in vitro*, were cut into small parts using a sterile razor blade and incubated at 21°C with 10 ml of buffer (2% (w/v) cellulose, 0.08% (w/v) macerozyme) for 3.5 hours. After incubation, the protoplast suspension was filtered, and the protoplasts were collected by centrifugation at 4°C, in a bench-top centrifuge (fixed angle rotor; Eppendorf centrifuge 5415 D) at 700 rpm ($450 \times g$) for 8 minutes. The protoplast pellet was resuspended in 2 ml of 200 mM CaCl₂ to wash the protoplasts, and then they were centrifuged at 4°C, in a bench-top centrifuge (fixed angle rotor; Eppendorf centrifuge 5415 D) at 600 rpm ($330 \times g$) for 5 minutes. After centrifugation, the protoplasts were resuspended in 2 ml of W5 solution (0.1% (w/v) glucose, 0.08% (w/v) KCl, 0.9% (w/v) NaCl, and 1.84% (w/v) CaCl₂, pH 5.8). Plasmid DNA samples (5 µg) were transferred to protoplasts (2×10^6), then 110 µl of PEG 4000 solution was added carefully from the side of the tube with a cut yellow tip (PEG solution is prepared fresh and cooled to room temperature before usage). Then it was mixed thoroughly by rotating tube and incubated in the dark for 16-24 hours before analysis. The expression of the YFP constructs in protoplasts was analyzed using a fluorescence microscope (Nikon TE-2000E) equipped with filters for analysing YFP (exciter HQ500/20x, emitter HQ535/30m) and chlorophyll autofluorescence (exciter D480/30x, emitter D660/50m).

Chapter 3

Molecular Genetic Study of Tic20 Homologues in *Arabidopsis thaliana*

3.1. Abstract

In this study *Arabidopsis thaliana* has been used as a plant model to investigate the involvement of Tic20 (translocon at the inner envelope membrane of chloroplasts, 20 kD) in protein import. Tic20 was originally identified biochemically in pea as a chloroplast protein import apparatus component. In *Arabidopsis*, there are four Tic20 homologues called (*atTIC20-I*, *atTIC20-IV*, *atTIC20-II*, *atTIC20-V*), all with predicted topological similarity to the pea protein *Pisum sativum* Tic20 (psTic20). Tic20 is an integral protein of the chloroplast inner envelope membrane. The topology of Tic20 with four putative α -helical domains and the cross-linking of preproteins to Tic20 (which increases at the later stages protein import) led to the assumption that Tic20 functions as a protein conducting component of the TIC complex, either independently or in conjunction with a third inner membrane component, Tic110. I used, TargetP analysis to predict the subcellular localization of all four *Arabidopsis* Tic20 proteins to the chloroplast. Moreover, the TMHMM program was used to identify the transmembrane domains of the atTic20 proteins in comparison to psTic20, to reveal any topological similarities that these proteins might have; the results showed that all atTic20 homologues have four α -helices. To test the TargetP predictions, envelope localization of each one was confirmed by transient expression of YFP fusions in *Arabidopsis thaliana* protoplasts. Furthermore, quantitative RT-PCR (and public microarray data from Genevestigator) revealed that the all atTIC20 homologues genes are expressed throughout development; *atTIC20-I* expression was highest in photosynthetic tissues, whereas *atTIC20-IV* expression was strong in non-photosynthetic tissues and seeds. To assess functional significance of the genes *in vivo*, T-DNA mutants were identified. Homozygous *tic20-I-1* and *tic20-I-2* plants have an albino phenotype which correlated with abnormal chloroplast development. By contrast, knockouts for the other three genes were indistinguishable from wild type. To test for functional redundancy, various double and triple mutants were studied; apart from those involving *tic20-I*, these were all phenotypically similar to wild type. The *tic20-I tic20-II* and *tic20-I tic20-V* doubles were albino, like the corresponding *tic20-I* parent. In contrast, *tic20-I tic20-IV* double homozygotes could not be identified, due to gametophytic and embryonic lethality. Redundancy between atTic20-I and atTic20-IV was confirmed by the partial complementation of *tic20-I* using an *atTIC20-IV* overexpressor construct. We conclude that atTic20-I and atTic20-IV are the major

functional Tic20 isoforms in *Arabidopsis*, with partially overlapping roles. While the atTic20-II and atTic20-V (Group 2 proteins) may not be required for normal *Arabidopsis* development under standard conditions.

Division of labour. The contributors to this Chapter were as follows: Ali Reza Kasmati, Ramesh Patel, Sybille Kubis, Jocelyn Bédard, Mats Töpel, Ghulam Murtaza and Paul Jarvis. Ramesh Patel assisted me with the identification and characterization of the T-DNA insertion lines. Sybille Kubis assisted me to generate the atTic20-I, atTic20-IV, atTic20-II and atTic20-V YFP fusion plasmids, and with the transfection and fluorescence microscopy studies. Jocelyn Bédard helped me with cDNA and T-DNA sequences. Mats Töpel assisted me with the phylogenetic analyses. Ghulam Murtaza and Ramesh Patel helped me with analysis of the atTic20 overexpression lines. All other experiments were carried out by Ali Reza Kasmati. This research was conducted under the supervision of Paul Jarvis.

3.2. Introduction

As I summarized in Chapter 1, the vast majority of chloroplast proteins are translated on cytosolic ribosomes and subsequently imported into plastids (Soll and Schleiff, 2004; Kessler and Schnell, 2006b; Smith, 2006; Inaba and Schnell, 2008; Jarvis, 2008b). Most of these proteins are synthesized as soluble precursors with N-terminal extensions called transit peptides and post-translationally imported into the plastid after binding to the chloroplast outer envelope membrane. The cleavable transit peptide is essential for chloroplast targeting and protein translocation across the envelope membranes, through complexes termed TOC and TIC (Translocon at the outer/inner envelope membrane of chloroplasts).

Biochemical studies using pea chloroplasts have resulted in the identification of several TOC and TIC components. The core TOC system consists of preprotein receptors (Toc159 and Toc34) together with a transmembrane channel protein (Toc75). While several TIC components have also been identified (Tic110, Tic62, Tic55, Tic40, Tic32, Tic22, Tic21 and Tic20), the exact roles of these proteins in the import process are not well understood (Soll and Schleiff, 2004; Kessler and Schnell, 2006b; Smith, 2006; Inaba and Schnell, 2008; Jarvis, 2008b). Arguably, the most crucial function of the TIC complex is the channel formation. However, there is a uncertainty about the identity of the relevant component(s), as it has been suggested that Tic20 (Kouranov *et al.*, 1998), Tic110 (Heins *et al.*, 2002; Balsera *et al.*, 2009), and Tic21 (also called Chloroplast Import Apparatus 5) (Teng *et al.*, 2006) may each perform this role. The Tic20 and Tic21 proteins share topological similarity and complementary expression patterns, and so it was initially proposed that the former operates during early plant development with the latter taking over later on (Teng *et al.*, 2006). Nonetheless, both proteins have been detected together in a 1 MDa complex, and it is suggested that this corresponds to the core, channel-forming TIC complex (Kikuchi *et al.*, 2009). Interestingly, Tic110 was absent from this assembly. Instead, it is suggested that Tic110 acts later in the translocation process by serving as a scaffold for the coordination of stromal chaperones (Inaba *et al.*, 2003) that bind to preproteins as they emerge from the import machinery (Inaba *et al.*, 2003). Such chaperones may drive translocation and/or facilitate protein folding following transit peptide cleavage.

Since little is known about the TIC complex, other transport machinery might give clues to the mechanism of plastid protein import. Phylogenetic studies suggest that

Tic20 is distantly related to bacterial branched amino acid transporters and the mitochondrial protein import channel protein, Tim17 and Tim23 (Rassow *et al.*, 1999; Reumann *et al.*, 2005). The putative homology and topological similarity of Tic20 with mitochondrial Tim17/22/23 preprotein translocase components supports the idea that Tic20 is an import channel (Reumann *et al.*, 1999, 2005), the notion of a channel-forming role for Tic20 was further supported. Nonetheless, there is a lack of *in vitro* evidence to substantiate this proposal (Soll and Schleiff, 2004). In mitochondria there are two TIM complexes; TIM23 and TIM22. The *Saccharomyces cerevisiae* (yeast) Tim23 (ScTim23) protein is the main part of TIM23 complex and it forms the channel, which mediates the translocation of proteins carrying an N-terminal presequence (Truscott *et al.*, 2001). These data suggested that Tic20 functions as a component of the protein-conducting channel at the inner envelope membrane, but so far there are no *in vitro* data available to support this idea.

The pea Tic20 (psTic20) protein was identified by chemical crosslinking and shown to be closely positioned to precursor proteins engaged in import at an intermediate stage (Ma *et al.*, 1996; Kouranov and Schnell, 1997). It was initially proposed that Tic20 has three transmembrane α -helices, but it now seems clear that there are four such domains (Kouranov *et al.*, 1998; Chen *et al.*, 2002). The predicted topology of Tic20 makes it a good candidate for the protein-conducting channel of the inner membrane translocon. Current TMHMM study (this report) shows that the pea and atTic20 proteins likely all have four putative transmembrane α -helices. The predicted topology of Tic20 makes it a good candidate for a component of the protein-conducting apparatus of inner membrane translocon.

In the *Arabidopsis* genome four different Tic20-related sequences have been reported to exist, and these are named according to their chromosomal locations: atTic20-I, atTic20-IV, atTic20-II and atTic20-V (Bédard and Jarvis, 2005; Kalanon and McFadden, 2008). Amino acid sequence identity shared between these four sequences and psTic20 was reported to be 62%, 35%, 25% and 25%, respectively, across the aligning regions (Bédard and Jarvis, 2005). Thus, while it seems likely that atTic20-I is the direct functional orthologue of psTic20, roles in protein import for the other Tic20 homologues cannot be excluded. Antisense technology was employed to elucidate the role of atTic20-I in protein import *in vivo* (Chen *et al.*, 2002). The antisense plants exhibited a pale-green phenotype, growth defects, small and underdeveloped chloroplasts, and reduced levels of chloroplast proteins; all this was associated with

reduced atTic20-I protein levels. Protein import studies using chloroplasts isolated from atTic20-I antisense plants revealed that they are defective in preprotein translocation at the level of the inner envelope membrane (Chen *et al.*, 2002). More recently, complete knockout of atTic20-I was reported to cause albinism (Teng *et al.*, 2006), and this was correlated with an import defect with apparent specificity for photosynthesis-related preproteins (Kikuchi *et al.*, 2009).

Recently, the completed genome of *Chlamydomonas reinhardtii* Tic20 (CrTic20) (Merchant *et al.*, 2007) provided an opportunity to predict which components are used for chloroplast protein translocation complexes. In comparison to atTic20, CrTic20 contained nearly all the components found in *Arabidopsis thaliana*. However, atTic20 has four homologues whereas CrTic20 has only two homologues. The CrTic20 study predicted that one of these homologues is more similar to atTic20-I, whereas the second homologue is more similar to atTic20-II and atTic20-V and may or may not be involved in protein translocation (Kalanon and McFadden, 2008). Like psTic20 and atTic20, the CrTic20 also contains four predicted transmembrane α -helices (like atTic20-I, the first helix of CrTic20 is also weakly predicted), and a predicted chloroplast-targeted transit peptide (Kalanon and McFadden, 2008).

Another member of a highly diverged family of the Tic20 protein was identified in *Apicomplexa*. Like other Tic20 proteins (pea and *Arabidopsis*) the *Toxoplasma gondii* Tic20 (TgTic20) is found to be an integral protein of the plastid inner membrane. *In silico* study of this protein revealed that it also contains four transmembrane domains, which are close at the C terminus end of this protein (van Dooren *et al.*, 2008). It was shown that the N-terminal portion of this protein is cleaved upon import to make the mature form of this protein with 20 kDa.

In this study, we analysed all four *Arabidopsis* Tic20-related sequences with the aim of revealing their phylogenetic and functional interrelationships.

3.3. Results

3.3.1 Bioinformatics

3.3.1.1. *Arabidopsis thaliana* Tic20 Homologue Protein

Alignments

The aim of this experiment was to determine the similarity between psTic20 and the four atTic20 homologues. According to our database searches (as stated previously), in *Arabidopsis* there are four homologues (atTic20-I, atTic20-IV, atTic20-II and atTic20-V) these may function in a similar style to psTic20 in protein import pathways. The identification of full-length, sequenced cDNA clones for each gene (accession numbers AK117165, AF361633, AY050346 and BT001996, respectively), confirmed that they are all expressed in plants. Protein sequences predicted using these cDNA sequences were analysed *in silico* experiments. Moreover, analysis using the TMHMM server indicated that the four proteins are topologically similar to psTic20 (Krogh *et al.*, 2001); *i.e.*, they each have four predicted transmembrane domains, and these are similarly located (Figure 3.1a). Analysis of the sequence alignment reveals high sequence similarity of psTic20 to atTic20-I (>62% identity), whereas atTic20-V shares much less similarity to psTic20 (19.0%), as shown in Figure 3.1a. The atTic20 homologues have been compared to other species such as the rice homologues (*Oryza sativa*) proteins, osTic20 (XP-507375), osTic20 (XP-466721) (Figure 3.1b).

(a)

1	2	3	4	5	
	62.8	29.4	23.5	19.0	1
		33.3	22.0	15.4	2
			20.9	17.7	3
				37.9	4
					5

psTic20 protein
atTic20-I protein (AK117165)
atTic20-IV protein (AF361633)
atTic20-II protein (AY053346)
atTic20-V protein (BT001996)

(b)

1	2	3	4	5	6	
	65.0	36.5	28.6	19.5	16.7	1
		34.3	28.8	17.9	15.2	2
			34.4	23.8	16.6	3
				16.7	15.5	4
					37.9	5
						6

osTic20 protein (XP466721)
osTic20 protein (XP507375)
atTic20-I protein (AK117165)
atTic20-IV protein (BX828045)
atTic20-II protein (BX821678)
atTic20-V protein (AY087311)

Figure 3.1. (a). The psTic20 and atTic20 protein sequences comparison. Percent amino acid identities shared between psTic20 and the atTic20 homologues. Values were calculated using MegAlign (ClustalW).

(b) Percent amino acid identities shared between the atTic20 and osTic20 homologues. Values were calculated using MegAlign (ClustalW).

3.3.1.2. Comparison with Possibly Related Mitochondrial Proteins and Phylogenetic Analysis

The aim of this experiment was to compare atTic20 homologues to mitochondrial proteins (atTim17-22-23), and assess whether the atTic20 homologues are truly homologous to atTim families. The amino acid sequences of Tic20 and Tim homologues from many species (Table 3.1) were retrieved by BLAST searching. Although previous study indicating a phylogenetic relationship between Tic20, bacterial branched-chain amino acid transporters, and the mitochondrial Tim17-22-23 proteins (Rassow *et al.*, 1999; Reumann *et al.*, 1999), we could find no such link. Attempts were made to align the Tic20 dataset with these other proteins, to include them in our phylogenetic analyses, but this was found to be impossible as homologous characters in the different datasets could not be identified. However, despite the apparent lack of homology, we tested the proposed similarity as follows. First, we manually aligned a representative set of Tic20 and Tim17-22-23 proteins to the prokaryotic sequences, and adjusted the alignment according to the one presented by Rassow *et al.* (1999). We then created four consensus sequences, one for each of the following groups: atTic20/psTic20, atTim17-22-23, yeast (*Saccharomyces cerevisiae*) scTim17/23 and the prokaryote proteins (Figure 3.2). Consensus sequences are useful when comparing distantly-related proteins as they only show conserved positions less prone to accumulate mutations; *i.e.*, functional sites or positions required for the maintenance of the three-dimensional structure (Lesk and Fordham, 1996). Our analysis shows that no position in the four consensus sequences shares the same residue, and that the Tic20 sequence does not match with the prokaryotic sequence in any position. Therefore, we concluded that the similarity previously reported between Tic20, the prokaryotic branched-chain amino acid transporters, and Tim17-22-23 is a result of convergent evolution and not evidence of homology (Figure 3.2).

Additionally, Figure 3.3 shows that the atTic20 family diverged into two distinct sub-clades (atTic20-IV/atTic20-I, and atTic20-II/atTic20-V), termed Group 1 and Group 2. The results suggest that there are two sub-types of atTic20-related protein in plants which diverged very early during the evolution of plants, but after primary endosymbiotic event. From this study it could be concluded that the all different atTic20 homologues, in pea and other species, do share conserved sequences.

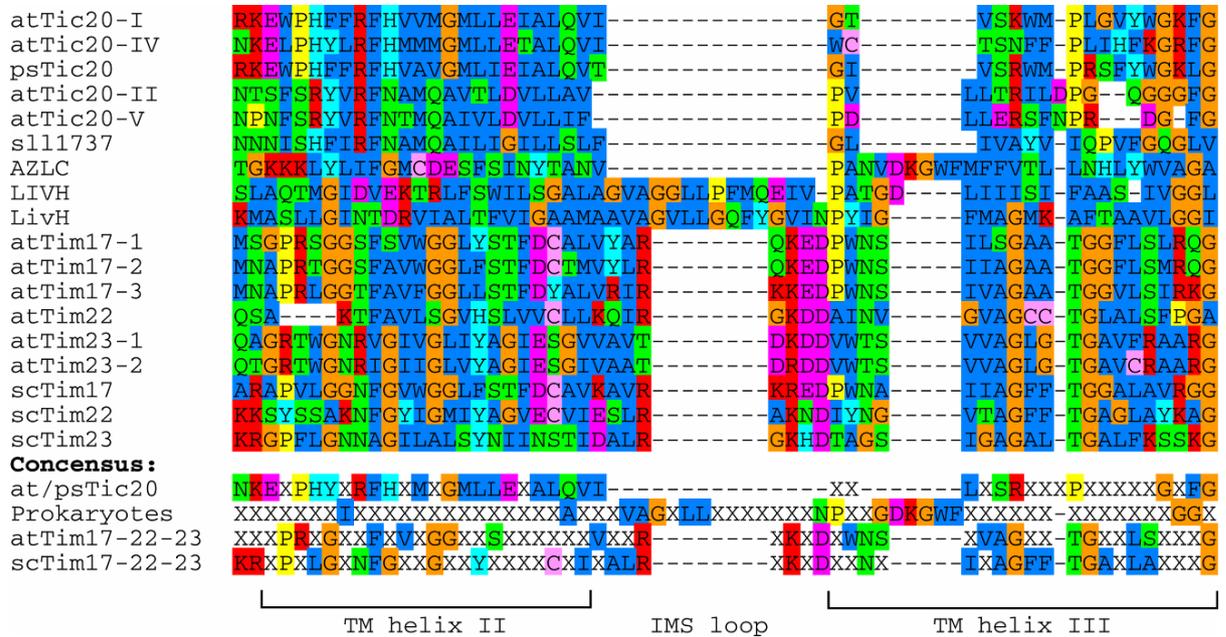


Figure 3.2. Assessment for possible homology between Tic20, bacterial amino acid transporter, and Tim17-22-23 sequences.

Representative sets of Tic20 proteins, prokaryotic branched-chain amino acid transporters, and Tim17-22-23 sequences from plants and yeast were aligned by mafft-linsi (Kato *et al.*, 2005) and manually adjusted to be comparable to the alignment presented by Rassow *et al.* (1999) (upper panel), and used to produce four consensus sequences. These consensus sequences revealed a lack of similarity between the different groups of proteins (lower panel); the residues shown are the ones found in 60% or more of the sequences, and the indicated transmembrane (TM) and intermembrane space (IMS) regions are based on Rassow *et al.* (1999). Thus, similarity previously reported between these different proteins is concluded to be a result of convergent evolution and not evidence of homology. The four groups are represented by the following proteins: Tic20 sequences from *Arabidopsis thaliana* and *Pisum sativum* (accession numbers: NP_171986 [atTic20-I], NP_192241 [atTic20-IV], AAC64607 [psTic20], NP_566112 [atTic20-II], AAP13363 [atTic20-V]); prokaryotic sequences (accession numbers: BAA17427 [*Synechocystis* sp., sll1737], CAA71940 [*Bacillus subtilis*, AZLC], Q58665 [*Methanocaldococcus jannaschii*, mjLivH], CAQ33776 [*Escherichia coli*, ecLivH]); Tim17-22-23 sequences from *Arabidopsis thaliana* (accession numbers: AAO63303 [atTim17-1], NP_973621 [atTim17-2], NP_196730 [atTim17-3], ABD64058 [atTim22], NP_564028 [atTim23-1], NP_177419 [atTim23-2], NP_187131 [atTim23-3]) (nomenclature and selection of atTim proteins follows Murcha *et al.* 2007); and, Tim17-22-23 sequences from *Saccharomyces cerevisiae* (accession numbers: CAA89438 [scTim17], CAA98795 [scTim22], CAA96296 [scTim23]).

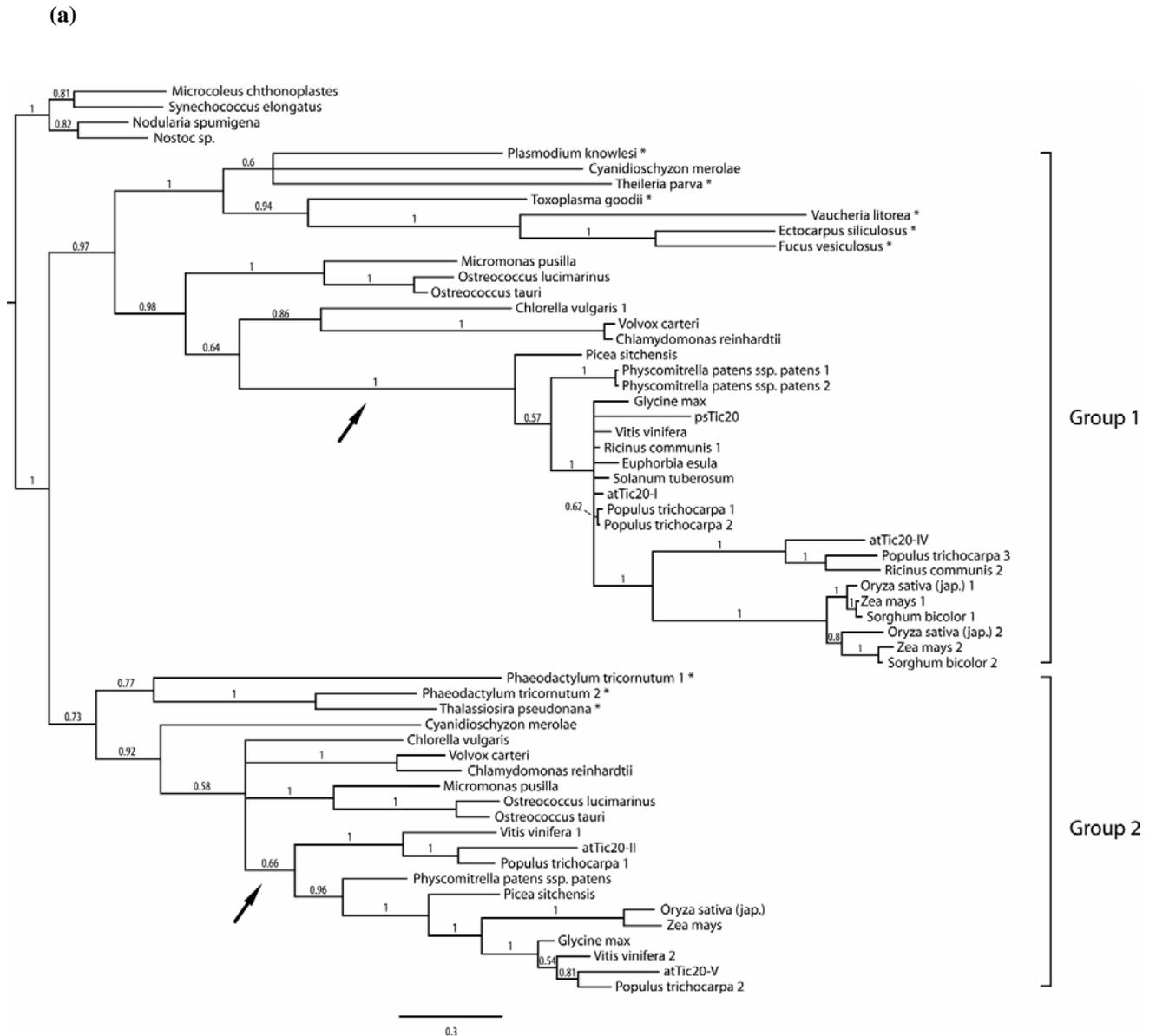


Figure 3.3. Phylogenetic analysis of Tic20-related sequences from different species.

Amino acid sequences of the TargetP-predicted mature regions of the indicated Tic20 homologues were aligned and used to produce a phylogenetic tree. The length of each branch in the phylogram is proportional to the number of expected changes along each branch (see scale). Posterior probability values are given above the branches. The two major clades are termed Group 1 and Group 2, as indicated. Gene and accession numbers of the sequences used are provided in Table 3.2. Species of origin is indicated with the following prefixes: ps, *Pisum sativum*; at, *Arabidopsis thaliana*; mt, *Medicago truncatula*; os, *Oryza sativa* (japonica subspecies); osi, *Oryza sativa* (indica subspecies); pp, *Physcomitrella patens*.

Species	Group	Name	NCBI accession nr.
<i>Microcoleus chthonoplastes</i>	Cyanobacteria		ZP_05027484
<i>Nodularia spumigena</i>	Cyanobacteria		ZP_01632509
<i>Nostoc sp.</i>	Cyanobacteria		NP_488844
<i>Synechococcus elongatus</i>	Cyanobacteria		YP_172029
<i>Arabidopsis thaliana</i>	1	atTic20-I	NP_171986
<i>Arabidopsis thaliana</i>	1	atTic20-IV	NP_192241
<i>Chlamydomonas reinhardtii</i>	1		
<i>Chlorella vulgaris 1</i>	1		
<i>Cyanidioschyzon merolae</i>	1		NP_848999
<i>Ectocarpus siliculosus</i>	1		CAT18806
<i>Euphorbia esula</i>	1		AAF34764
<i>Fucus vesiculosus</i>	1		CAX12421
<i>Glycine max</i>	1		ACU17555
<i>Micromonas pusilla</i>	1		
<i>Oryza sativa (jap.) 1</i>	1		NP_001060030
<i>Oryza sativa (jap.) 2</i>	1		NP_001047425
<i>Ostreococcus lucimarinus</i>	1		
<i>Ostreococcus tauri</i>	1		
<i>Physcomitrella patens ssp. patens 1</i>	1		XP_001775811
<i>Physcomitrella patens ssp. patens 2</i>	1		XP_001762298
<i>Picea sitchensis</i>	1		ABK27145
<i>Pisum sativum</i>	1	psTic20	AAC64607
<i>Plasmodium knowlesi</i>	1		XP_002259450
<i>Populus trichocarpa 1</i>	1		XP_002307678
<i>Populus trichocarpa 2</i>	1		XP_002300750
<i>Populus trichocarpa 3</i>	1		XP_002329357
<i>Ricinus communis 1</i>	1		XP_002510436
<i>Ricinus communis 2</i>	1		XP_002517985
<i>Solanum tuberosum</i>	1		ABB86268
<i>Sorghum bicolor 1</i>	1		XP_002463046
<i>Sorghum bicolor 2</i>	1		XP_002452444
<i>Theileria parva</i>	1		XP_765367
<i>Toxoplasma goodii</i>	1		EEE19444
<i>Vaucheria litorea</i>	1		YP_002327553
<i>Vitis vinifera</i>	1		XP_002281211
<i>Volvox carteri</i>	1		
<i>Zea mays 1</i>	1		NP_001140974
<i>Zea mays 2</i>	1		NP_001149306
<i>Arabidopsis thaliana</i>	2	atTic20-II	NP_566112
<i>Arabidopsis thaliana</i>	2	atTic20-V	AAP13363
<i>Chlamydomonas reinhardtii</i>	2		
<i>Chlorella vulgaris</i>	2		
<i>Cyanidioschyzon merolae</i>	2		
<i>Glycine max</i>	2		ACU20747
<i>Micromonas pusilla</i>	2		EEH57093
<i>Oryza sativa (jap.)</i>	2		EEE57727
<i>Ostreococcus lucimarinus</i>	2		XP_001419447
<i>Ostreococcus tauri</i>	2		
<i>Phaeodactylum tricorutum 1</i>	2		XP_002176871
<i>Phaeodactylum tricorutum 2</i>	2		XP_002179527
<i>Physcomitrella patens ssp. patens</i>	2		XP_001752651
<i>Picea sitchensis</i>	2		ABK25967

<i>Populus trichocarpa 1</i>	2	XP_002328962
<i>Populus trichocarpa 2</i>	2	XP_002328198
<i>Thalassiosira pseudonana</i>	2	XP_002287042
<i>Vitis vinifera 1</i>	2	CAO22941
<i>Vitis vinifera 2</i>	2	XP_002267869
<i>Volvox carteri</i>	2	
<i>Zea mays</i>	2	ACF84989

Table 3.1. The organism abbreviation, the Tic20-related amino acid sequences used for the phylogenetic analysis.

3.3.1.3. Prediction of Protein Localization (by TargetP)

The aim of this experiment was to develop a better understanding of subcellular localization of atTic20 homologues. Subcellular localization of proteins is essential for the structure and function of the cell. Consequently, awareness of the subcellular localization of proteins is crucial to recognize their roles and interacting partners in cellular metabolism (Kumar *et al.*, 2002; Huh *et al.*, 2003). The TargetP program predicted that all four *Arabidopsis* Tic20 homologues have a transit peptide with high confidence (Emanuelsson *et al.*, 2000), which is consistent with localization to the inner envelope membrane (Hofmann and Theg, 2005b). The results are shown in Table 3.2 (all raw data for this experiment are attached in Appendix 3.1).

Amino acid sequences of psTic20 and four atTic20 homologues were chosen for the TargetP analyses. Predictions for chloroplast localization were achieved by using this web-site (<http://www.cbs.dtu.dk/services/TargetP/>). The TargetP prediction provides a Reliability Class (RC) score according to the difference between the highest and the second-highest network output score; this feature specifies the level of certainty in the prediction for a particular sequence. The lower the RC, the better the prediction which is more reliable (Table 3.1). The results shown in Table 3.2 illustrate that the prediction results using TargetP method are reliable.

Protein	Predicted Location	Reliability Class	TP Length (aa)
psTic20	Chloroplast	4	42
atTic20-I	Chloroplast	4	65
atTic20-IV	Chloroplast	4	48
atTic20-II	Chloroplast	3	49
atTic20-V	Chloroplast	4	49

Table 3. 2. TargetP was used to predict the location of each protein; the Reliability Class and the predicted length of the transit peptide (TP) are also shown.

3.3.1.4. Prediction of Transmembrane Domains (by TMHMM)

The aim of this experiment was to predict the transmembrane (TM) domains of the four atTic20 proteins in comparison to psTic20, to identify any topological similarities that they may have. Analysis using the TMHMM server indicated that the four proteins are topologically similar to psTic20 (Krogh *et al.*, 2001) (Figure 3.3 and Table 3.3). The same amino acids sequences were used in this experiment as were used previously in the TargetP section to produce these results. Predictions for TM domains were achieved by using this web-site (<http://www.cbs.dtu.dk/services/TMHMM/>), and used to create Table 3.3 and Figure 3.4. As shown in Figure 3.4, all four atTic20 homologues have very similar TM domains compared to psTic20; i.e., they all have four predicted TM domains and these are similarly located. All TMHMM output files are attached in Appendix 3.2. As a result of these TM predictions, it may be further concluded that the putative atTic20 homologues are close relatives of psTic20. Based on these observations, I propose that the atTic20 homologues have a similar topology to psTic20.

	Transmembrane Helix 1 (aa)	Transmembrane Helix 2 (aa)	Transmembrane Helix 3 (aa)	Transmembrane Helix 4 (aa)
psTic20	101-120	140-162	174-196	211-233
atTic20-I	121-143	165-187	194-216	231-253
atTic20-IV	124-146	161-183	196-218	233-255
atTic20-II	61-83	98-120	132-154	169-191
atTic20-V	62-86	106-123	130-152	172-194

Table 3.3. This presents the list of transmembrane domains within psTic20 and the atTic20 homologues, as predicted by the program TMHMM.

3.3.2. Protein Localization by Yellow Fluorescent Protein (YFP)

The aim of the present work was to provide experimental support for the TargetP predictions. The subcellular localization of the *Arabidopsis* Tic20 homologues was assessed by the analysis of YFP fusion proteins. To this end, full-length cDNA sequences (Table 2.4) for each of the four *Arabidopsis* *TIC20* genes were inserted into the p2GWY7 vector by Gateway recombination cloning (Karimi *et al.*, 2005); this vector adds a C-terminal YFP tag (Appendix 3.3, Supplementary Figure 3.1). *Arabidopsis* protoplasts were then transfected using these constructs, and analysed by fluorescence microscopy (Figure 3.4). The red fluorescent signal of chlorophyll shows the location of the chloroplasts in each case. For all four fusion proteins, the yellow-green fluorescence of YFP was observed in a ring-like pattern around the periphery of each chloroplast. This pattern was strongly reminiscent of the distributions seen for atTic110:YFP and atTic40:YFP fusion proteins in a previous, experimentally very similar study (Bédard *et al.*, 2007), indicating that the *Arabidopsis* Tic20 homologues are all targeted to the chloroplast envelope. In view of the fact that most outer envelope membrane proteins do not possess a transit peptide (as the four *Arabidopsis* Tic20 homologues do) (Hofmann and Theg, 2005b), and considering the aforementioned topological similarities and phylogenetic relationship between the *Arabidopsis* proteins and psTic20, these results strongly support the hypothesis that *Arabidopsis* Tic20 homologues are localized in the chloroplast inner envelope membrane.

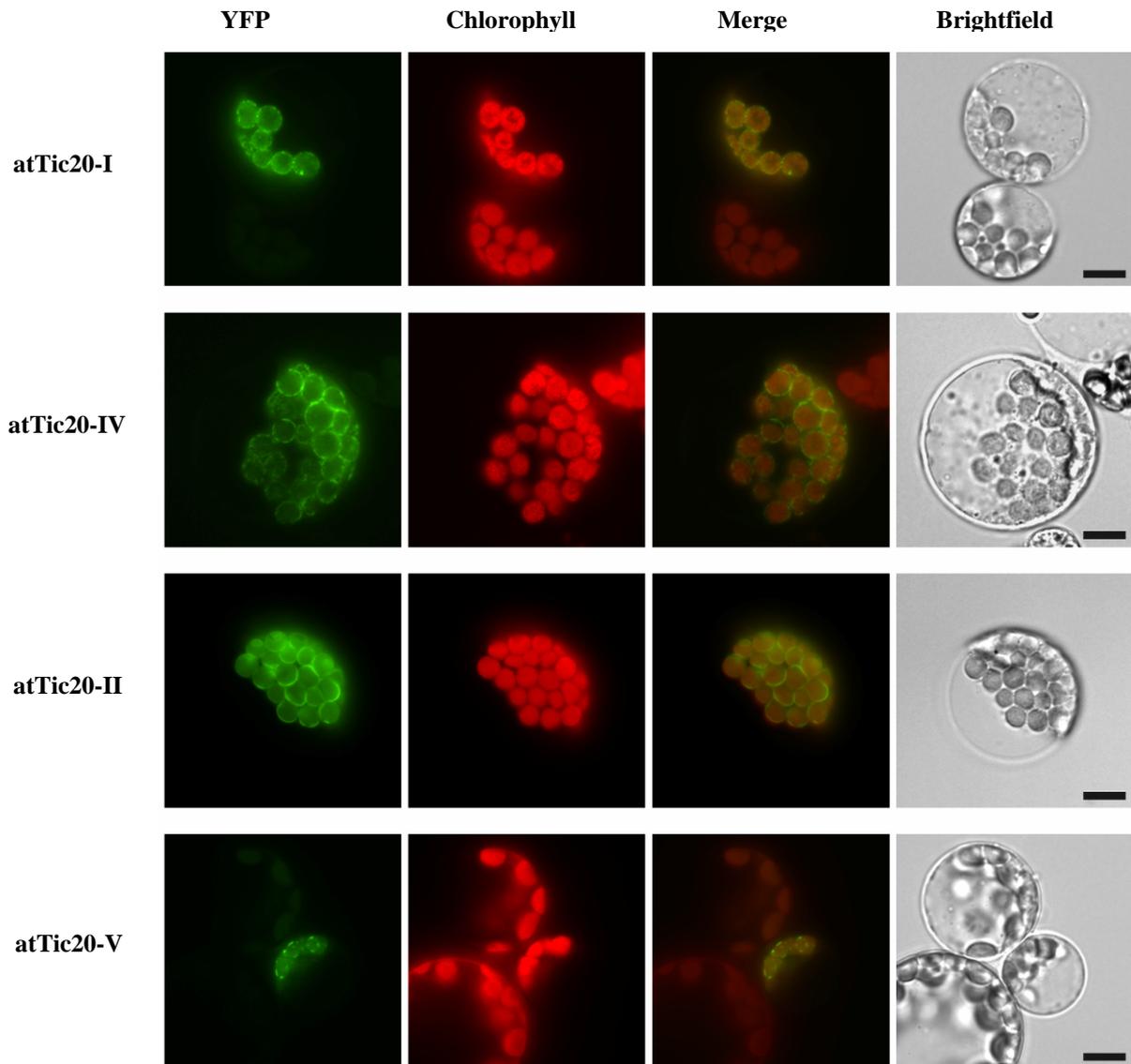


Figure 3.4. Subcellular localization of the *Arabidopsis* Tic20 proteins as assessed by YFP fusion-protein analysis.

Wild-type *Arabidopsis* protoplasts were transfected with the indicated plasmids (atTic20-I:YFP, atTic20-IV:YFP, atTic20-II:YFP and atTic20-V:YFP) and then analysed for YFP fluorescence (green, left panels) and chlorophyll autofluorescence (red, centre-left panels), as well as under bright field illumination (right panels). In addition, an overlay of the YFP and chlorophyll images is presented (centre-right panels). In all four cases, YFP fluorescence was exclusively associated with the chloroplast envelope. Scale bars = 10 μ m.

3.3.3. Genevestigator Analysis

To obtain information on *atTIC20* gene expression at different plant growth stages, publicly-available microarray data were analysed using the Genevestigator tool (Zimmermann *et al.*, 2004). A developmental time course revealed that three of the four *Arabidopsis TIC20* genes are expressed throughout the plant's life-cycle (Figure 3.5). Two TOC-related *Arabidopsis* genes (*atTOC33* and *atTOC34*), as well as the gene for a major component of the TIC complex, *atTIC110* were used in this experiment (Figure 3.5). As indicated in Figure 3.5, the control genes *atTOC33* and *atTIC110* were expressed most strongly, whereas *atTIC20-IV* was expressed at very low levels except in mature siliques where the expression level of *atTIC20-IV* was the same as *atTOC33*. In comparison to control genes, the *atTIC20-II* and *atTIC20-V* genes exhibited an intermediate expression level. Interestingly, in developed rosette young flower, and the developed flower the *atTIC20-V* expression was approximately similar to *atTIC110*. However, this shows, it is possible that this gene (*atTIC20-V*) play a role in developed rosette young flower, and the developed flower. In agreement with the RT-PCR data, *atTIC20-IV* exhibited relatively low levels of expression at photosynthetic developmental stages, and markedly elevated levels in mature siliques and germinating seeds; this supports the notion that *atTIC20-IV* is particularly important during seed development. A developmental time course revealed that three of the four *Arabidopsis Tic20* genes are expressed throughout the plant's life-cycle (Figure 3.5); unfortunately, *atTIC20-I* could not be analysed in this way due to the absence of a reliable probe set on the Affymetrix microarray (Teng *et al.*, 2006).

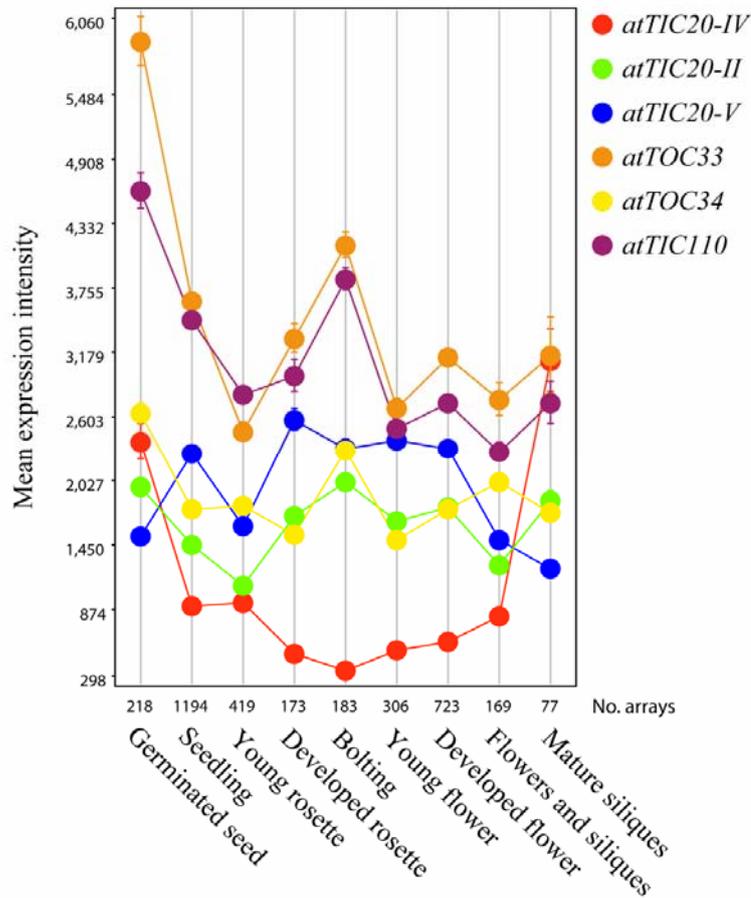


Figure 3.5. Expression of the *Arabidopsis* Tic20 genes using publicly-available microarray data. Affymetrix GeneChip data were analysed and retrieved using the Genevestigator V3 analysis tool (<https://www.genevestigator.ethz.ch/>) (Zimmermann *et al.*, 2004; Grennan, 2006). Presented data were prepared using the Meta-Profile Analysis tool using the Development representation in scatter-plot format. Data from all high-quality ATH1(22k) arrays were analysed; this amounted to a total of 3110 arrays. Values shown are means (\pm SE). The total number of arrays used to derive each data point shown is indicated. Data representations were exported from Genevestigator in Encapsulated PostScript format, and then compiled and annotated using appropriate graphics software. The genes analyzed were as follows: *atTIC20-IV* (At4g03320; red); *atTIC20-II* (At2g47840; green); *atTIC20-V* (At5g55710; blue); *atTOC33* (At1g02280; orange); *atTOC34* (At5g05000; yellow); and *atTIC110* (At1g06950; purple). The *atTIC20-I* gene could not be analysed due to the absence of a reliable probe set on the Affymetrix microarray (Teng *et al.*, 2006).

3.3.4. Expression Profiles of *Arabidopsis* TIC20 Homologues

To gain insight into the functional relationships between the *Arabidopsis* TIC20 homologues, their developmental and tissue-specific gene expression patterns were studied by quantitative real time RT-PCR (Figure 3.6). Samples analysed included rosette leaves, siliques and roots from mature plants. Additionally, we analysed whole 5-day-old seedlings grown in the dark (5dD) or in the light (5dL). The latter sample was compared to 14-day-old, light-grown seedlings (14dL) in order to study a developmental difference.

The result indicated that *atTIC20-I* expression is highest in young, actively growing photosynthetic tissues (5 dL and 14 dL; Figure 3.6), and that its expression declines in more mature tissues (Rosettes). By contrast, the expression of *atTIC20-I* was relatively weak in non-photosynthetic tissues (5 dD and Roots; Figure 3.6). These data suggest that *atTIC20-I* may be relatively more important for photosynthetic development, which is in-line with observations made in relation to the dominant isoforms of the TOC receptors (Bauer *et al.*, 2000; Kubis *et al.*, 2003). consistent with the Genevestigator study, this may suggested that *atTIC20-IV*, the other close homologue of psTic20 in *Arabidopsis*, may have a particularly important role in non-photosynthetic protein import.

The expression pattern of *atTIC20-IV* was distinctly different, and somewhat complementary to that of the other *atTIC20* homologues gene: *atTIC20-IV* expression levels were higher in the non-photosynthetic tissues than in light-grown seedlings and rosette leaves (Figure 3.6). In general, the *atTIC20-IV* gene was expressed at lower levels than *atTIC20-I*, indicating that *atTIC20-I* is the major isoform amongst the Group 1 proteins. In agreement with this assessment, when we conducted database searches using the BLAST program (Altschul *et al.*, 1990), we detected 90 expressed sequence tags (ESTs) for *atTIC20-I* compared to just 25 for *atTIC20-IV*. Nonetheless, expression of *atTIC20-IV* did exceed that of *atTIC20-I* in the non photosynthetic samples (5 dD and Roots; Figure 3.6). Finally, it is noteworthy that expression of *atTIC20-IV* was highest in siliques, suggesting an important role in seed development.

The other two genes (*atTIC20-II* and *atTIC20-V*) were expressed at relatively high levels throughout development, and shared rather similar patterns of expression (Figure 3.6); EST numbers detected as described above were 92 and 95, respectively. The expression profiles of *atTIC20-II* and *atTIC20-V* were both more similar to that of

atTIC20-I than to that of *atTIC20-IV* (amongst the Group 1 genes); nonetheless, the photosynthetic *vs.* non-photosynthetic expression intensity differential was considerably less for *atTIC20-II* than for *atTIC20-V*, and so in this sense the profile of the former was relatively more similar to that of *atTIC20-IV*. Generally, it appears that all four genes are expressed at different levels throughout development.

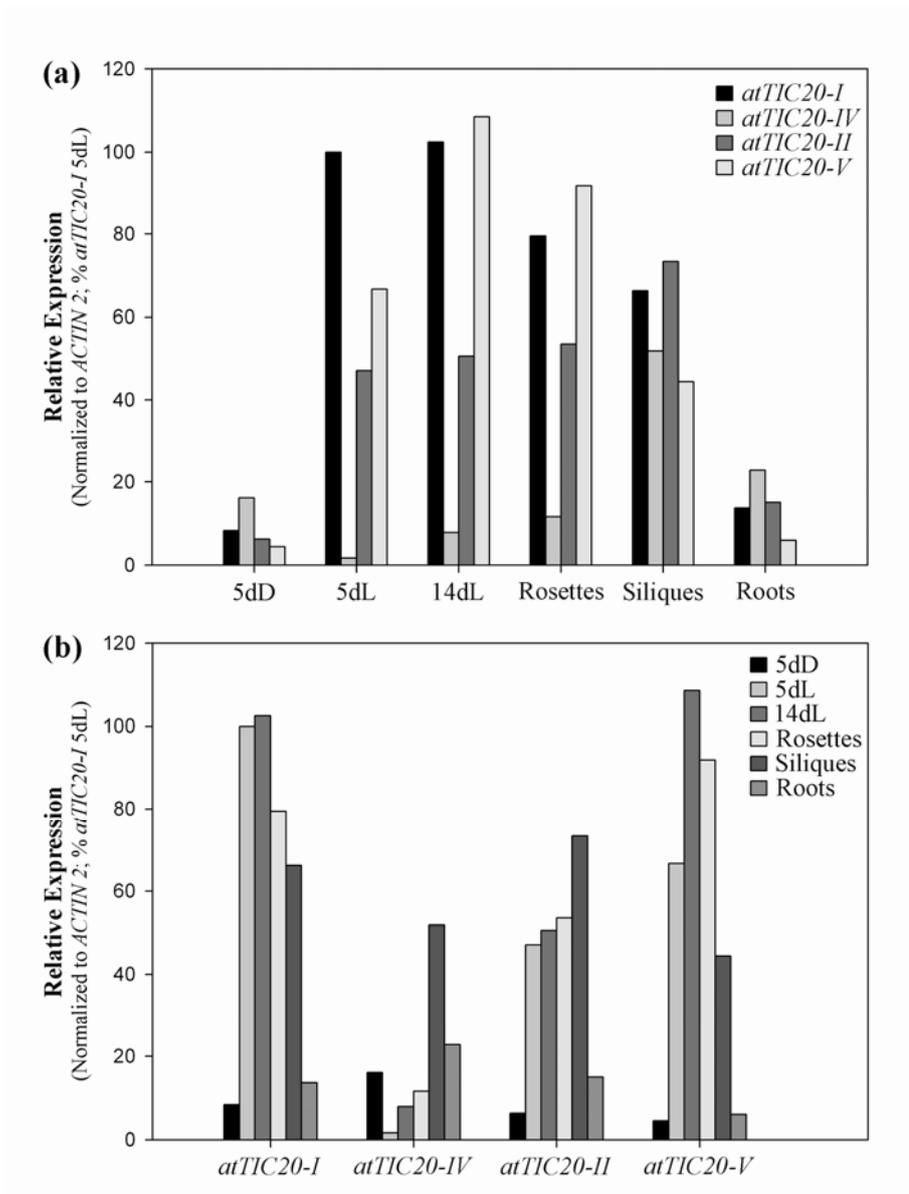


Figure 3.6. Quantitative RT-PCR analysis of the expression of the Tic20 genes in different tissues and at different stages of development.

Total RNA isolated from whole seedlings grown *in vitro* for five days in the dark (5dD), or five and 14 days in the light (5dL and 14dL, respectively), as well as from three different organs of 28-day-old plants grown on soil (roots, rosette leaves and floral buds). The RNA samples were representative of ~10-30 seedlings (5dD, 5dL and 14dL), or 5-25 mature plants (roots, rosettes and buds). Samples were analysed using quantitative real-time RT-PCR. Using the results, ΔC_T values were calculated for each Tic20 gene, relative to the control gene, *ACTIN 2* (At3g18780). These values were then expressed as percentages of the ΔC_T value for the most highly-expressed gene, *atTIC20-I* (5-day-old in the light). The data shown are means derived from three independent amplifications, and reflect the relative expression levels of the four genes. Panels (a) and (b) contain the same data presented in different ways, to aid interpretation.

3.3.5. Mutant Analysis of *Arabidopsis* *TIC20* Genes

In order to analyse the functional significance of the *Arabidopsis* *TIC20* homologues *in vivo*, two independent T-DNA insertion mutants for each of the four genes have been identified. Firstly, all of the T-DNA insertion sites were confirmed by genomic PCR, and by the sequencing of the T-DNA/gene junctions at one or both sides in each case, as indicated (Figure 3.7). Disruption of *atTIC20-I* causes an albino phenotype. However, none of the *tic20-IV*, *tic20-II* and *tic20-V* single mutants was detectably different from wild type. Therefore, double and triple-mutants combination were generated in order to investigate the possibility that this might be due to functional redundancy (see later).

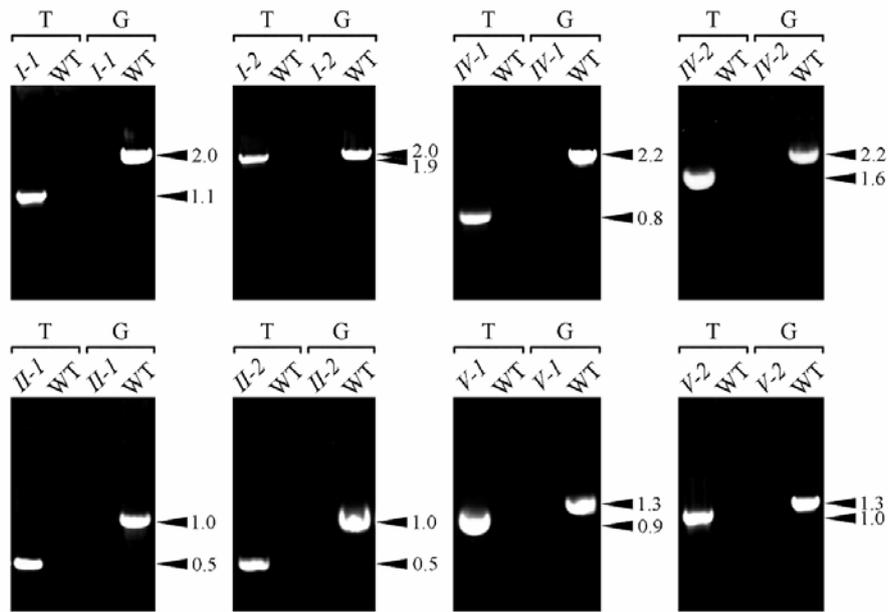


Figure 3.7. Analysis of the Tic20 T-DNA mutants by genomic PCR.

Genomic DNA samples extracted from wild type and putative homozygous mutants were analysed by PCR. Appropriate T-DNA- and *atTIC20*-gene-specific primers were employed. Two different primer combinations were used in each case: the first ('T') comprised one T-DNA border primer and one gene-specific primer (LB + reverse: *tic20-I-1*, *tic20-IV-1*, *tic20-IV-2*, *tic20-II-2*; LB + forward, *tic20-I-2*, *tic20-II-1*, *tic20-V-2*; RB + reverse, *tic20-V-1*); the second ('G') comprised two gene-specific primers flanking the T-DNA insertion site. The PCR products were resolved by agarose gel electrophoresis, and visualized by staining with SYBR Safe. Amplification using "T" indicated the presence of the mutant allele, whereas amplification using "G" indicated the presence of the wild-type allele; amplification with the former but not the latter demonstrated that the plant was homozygous mutant. The genotype names are shortened as follows: 'I-1' indicates *tic20-I-1*; 'I-2' indicates *tic20-I-2*; and so on. Sizes of the amplicons are indicated at right (in kb).

3.3.5.1. Segregation Analysis of *Arabidopsis* T-DNA Mutants

Segregation analysis was performed in order to ensure the identification of only single-locus insertion lines; a Mendelian ratio of three antibiotic-resistant plants to every one antibiotic-sensitive plant indicates the presence of just a single T-DNA insertion (Table 3.4). Further segregation analysis in later generations enabled the identification of homozygous lines for analysis, and the zygosity of these lines was confirmed by genomic PCR (Figure 3.7).

Initially, plating the *tic20-IV-1* mutant line on phosphinothricin-containing medium revealed a ratio close to 3:1, suggesting a single insertion for that family. However, homozygous plants were not found in the next generation. Therefore, it was presumed that this line was not genetically clean (perhaps there were unlinked, secondary mutations). Thus, that plant was back-crossed to wild type in order to clean the line. In the F3 generation, a homozygous line for *tic20-IV-1* was found. Apart from *tic20-I-1* and *tic20-I-2* (Figure 3.7) all knockout mutants in the homozygous state were found to show a wild-type-like phenotype. The genotype of all albino plants in the *tic20-I-1* and *tic20-I-2* lines were also confirmed by PCR and they were homozygous for the T-DNA insertion (Figure 3.7).

Parental Genotype	Parental Generation	Resistant (if Plated on selection)			Green : Bleached	Resistant : Sensitive	
		Selection	Green	Bleached			Sensitive
<i>+/tic20-I-1</i>	T6	None	2528	409	-	6.18 : 1.00	
<i>+/tic20-I-2</i>	T6	None	1937	609	-	3.18 : 1.00	
<i>+/tic20-I-1</i>	T5	Kanamycin	170	67	83	-	2.86 : 1.0
<i>+/tic20-I-2</i>	T5	Kanamycin	103	19	45	-	2.71 : 1.0
<i>tic20-IV-1</i>	F3	Phosphinothricin	292	0	97	-	3.01 : 1.0
<i>tic20-IV-2</i>	T2	Hygromycin	907	0	327	-	2.77 : 1.0
<i>tic20-II-1</i>	T4	Kanamycin	2705	0	810	-	3.08 : 1.0
<i>tic20-II-2</i>	T4	Sulfadizine	1776	0	562	-	3.16 : 1.0
<i>tic20-V-1</i>	T4	Hygromycin	274	0	96	-	2.85 : 1.0
<i>tic20-V-2</i>	T4	Kanamycin	692	0	250	-	2.77 : 1.0

Table 3.4. Genetic analyses of the atTic20 homologue mutants. Segregation analysis of the T-DNA-associated antibiotic resistance marker in each one of the Tic20 mutants.

3.3.5.2. Analysis of T-DNA Insertion Mutants by RT-PCR

To assess the effect of each T-DNA insertion on *TIC20* gene expression, RT-PCR analysis was conducted in each case (Figure 3.8b). The results confirmed that the relevant full-length mRNA was absent for all of the mutants except *tic20-I-2*, *tic20-II-1* and *tic20-II-2*; these three mutants displayed expression levels reduced to 29.0%, 8.4% and 58.0% of that in the wild type, respectively (Figure 3.8c), and so are considered to be knockdown alleles (Figure 3.8b). The other five mutants are all true knockout alleles that produced no mRNA. The mRNA samples for RT-PCR used for this experiment were derived from 10-day-old homozygous individuals grown *in vitro*. Amplification was performed using gene specific primers and products were visualized by staining with ethidium bromide following agarose gel electrophoresis. Additionally, the *atTOC33* and the *eIF4E1* genes were used as controls in each experiment. A “water” control (lacking template) was also included for all primer pairs, to verify the absence of contamination in the reagents.

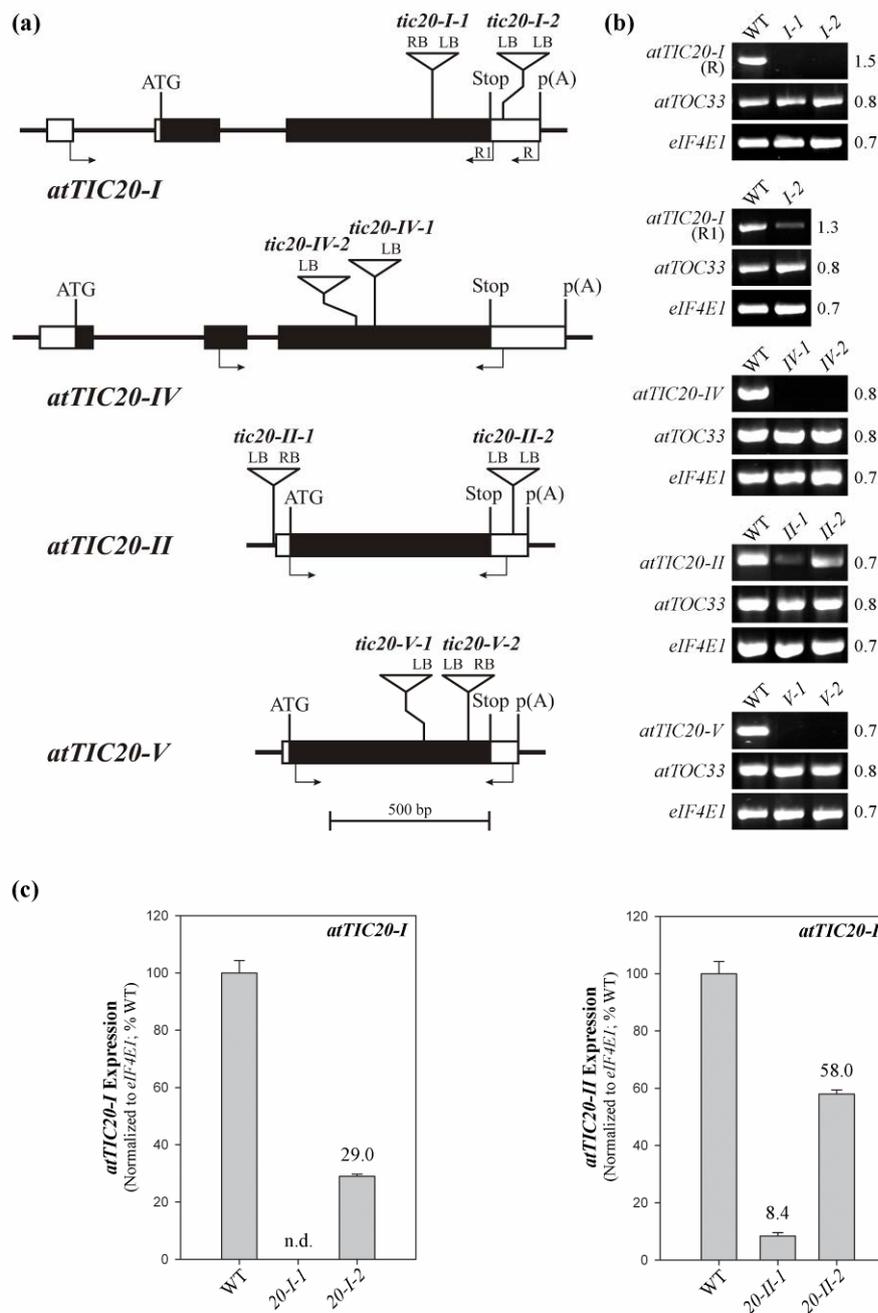


Figure 3.8. Molecular analysis of the Tic20 T-DNA insertion lines.

(a) Schematic diagrams showing the structure of each gene and the location of each T-DNA insertion. Protein-coding exons are represented by black boxes, and untranslated regions by white boxes; introns are represented by thin lines between the boxes. Locations of the primers used for RT-PCR analysis are shown by the arrows beneath each gene model. The T-DNA insertion sites are indicated precisely, but the insertion sizes are not to scale. ATG, translation initiation codon; Stop, translation termination codon; p(A), polyadenylation site; LB, T-DNA left border; RB, T-DNA right border.

(b and c) Analysis of the expression of the four *Arabidopsis* Tic20 genes in wild-type and mutant plants. The locations of the amplification primers used are shown in panel (a). Similar analysis of the expression of *atTOC33* and of the translation initiation factor gene, *eIF4E1* (At4g18040), was used to normalize loading. Sizes of the amplicons are indicated at right (in kb). The RNA samples were isolated from whole, 10-day-old homozygous plants grown *in vitro*, and were representative of ~20-30 seedlings. Because the data revealed that the *tic20-I-2*, *tic20-II-1* and *tic20-II-2* mutants were not null, semi-quantitative analysis was conducted in similar fashion (see charts in panel [c]); the values shown are means (\pm SE) derived from four independent amplifications, and are expressed as a percentage of the wild-type value for each gene. The data were normalized using similarly-derived data for *eIF4E1*. PCR amplification was performed over a total of 25 cycles.

3.3.5.3. Analysis of Albino Phenotype of *tic20-I* Mutants

Amongst the various mutant lines identified, only *tic20-I-1* and *tic20-I-2* exhibited a phenotype that was obviously different from wild type. In populations segregating for these two mutations, a significant number of individuals exhibited a striking albino-like appearance (Table 3.4). Co-segregation of the albino phenotypes with the T-DNA insertions was confirmed by PCR analysis; 25 albinos and 38 green plants from *tic20-I-1* populations, and 38 albinos and 65 greens from *tic20-I-2* populations were genotyped; all of the albinos were found to be homozygous mutant, whereas all of the greens were either heterozygous or wild type. The proportion of homozygous albino plants in such populations was consistently less than the 25% expected according to normal inheritance (particularly for the knockout allele, *tic20-I-1*), suggesting that the homozygotes have reduced viability at an early stage of development. Photographs of 38-day-old plants carrying these two mutations are shown in Figure 3.9; the atToc159 knockout mutant, *plastid protein import 2 (ppi2)*, is shown for comparison purposes, as this mutant also exhibits an albino phenotype (Bauer *et al.*, 2000). These plants were all grown on medium containing 3% sucrose, as they are unable to survive photoautotrophically. Interestingly, we observed that *tic20-I-2* plants are slightly larger than *tic20-I-1* plants, and that they are somewhat greener in appearance (Figure 3.9a); this is consistent with the observation that *tic20-I-2* is a knockdown allele, as described earlier (Figure 3.8b). The *ppi2* mutant grew to a larger size than either of the *tic20-I* mutants, which may be a reflection of the continued expression of related TOC receptor proteins in *ppi2* (Kubis *et al.*, 2004).

The block of greening in the *tic20-I* mutants suggested a defect in chloroplast development, and so plastid ultrastructure was analysed using transmission electron microscopy (Figure 3.10). The cotyledons of 10-day-old and 14-day-old (Figure 3.10), *in vitro* grown plants were analysed, and representative images are shown. In *tic20-I-1* plants, hardly any chloroplast development was observed. The *tic20-I-1* plastids did not contain any thylakoid membranes, and a significant proportion were found to contain large inclusions, or were surrounded by multilayered envelope membrane structures. By contrast, while the chloroplasts of the *tic20-I-2* mutant were much smaller than those in wild type, they did exhibit a significant amount of thylakoid membrane development, especially after 14 days of growth (Figure 3.10). This is consistent with the visibly greener appearance of plants carrying this allele (Figure 3.9a). Taken together, these

data support the notion that atTic20-I is essential for chloroplast biogenesis and development. Figure 3.10 also shows that the development of *ppi2* plastids is severely disturbed, as was reported previously (Bauer *et al.*, 2000; Kubis *et al.*, 2004), although the severity of the defect is less strong than that in *tic20-I-1* as the numerous inclusions and multilayered structures of the latter were not apparent.

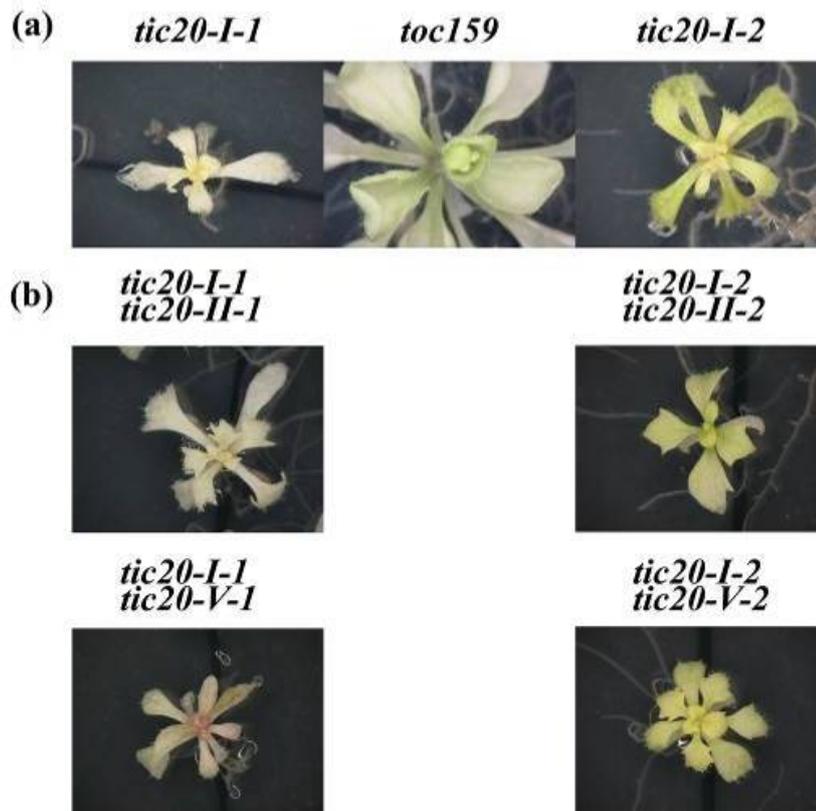


Figure 3.9. The albino phenotypes of *tic20-I* single- and double-mutant plants. Populations segregating for the indicated *tic20-I* mutation, either in the wild-type background (a) or in the indicated homozygous *tic20-II* or *tic20-V* mutant background (b), were plated on standard MS medium. After five days of growth, the homozygous albino mutants were rescued to medium containing 3% sucrose and grown side-by-side under standard conditions until they were 38 days old in total. The plants were then photographed, and representative images are shown. A similarly grown *ppi2* homozygous mutant plant is shown for comparison (a). Plants were illuminated with dim light ($\sim 10 \mu\text{mol}/\text{m}^2/\text{sec}$), under a standard long-day cycle, to aid growth.

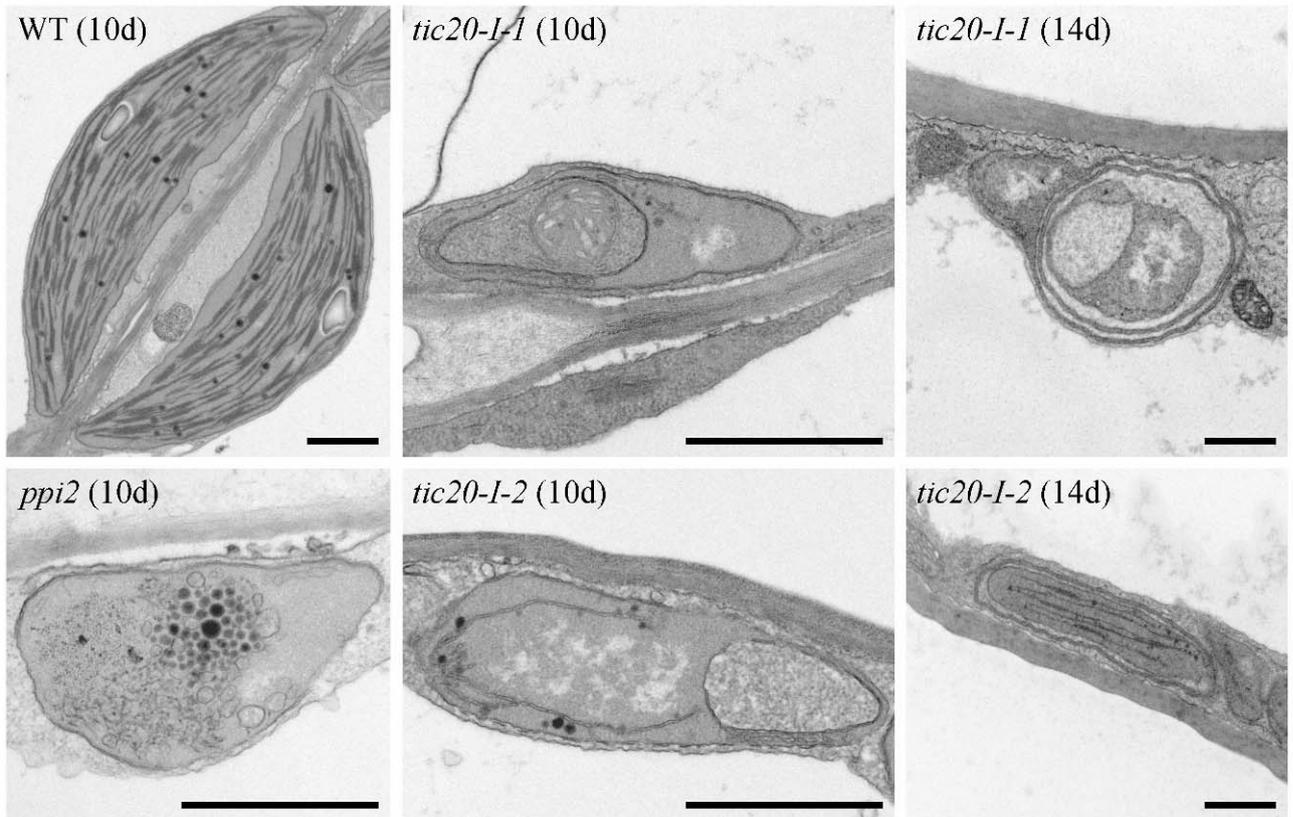


Figure 3.10. Analysis of plastid ultrastructure in the *tic20-I* single mutants.

The cotyledons of 10- and 14-day-old plants were analysed by transmission electron microscopy ('10d' and '14d' indicate different plant ages). Seeds were plated on standard MS medium; after five days of growth, the homozygous albino mutants were rescued to medium containing 3% sucrose and thereafter were grown side-by-side under standard conditions. On average, ~30 whole-chloroplast micrographs from each of three independent plants per genotype per age (a minimum of 55 chloroplasts per genotype per age) were analysed carefully, and used to select the representative images shown. Size bars = 1.0 μ m.

3.3.5.4. The Analysis of *tic20-I* Double Mutations

To study functional relationships between atTic20-I and the other three genes, *tic20-I* heterozygotes (both alleles) were crossed to the other mutants. Individual green plants from resulting F₂ (or F₃, in the case of *tic20-I-1 tic20-V-1*) generations were genotyped, and plants that were heterozygous for *tic20-I* and homozygous for each of the other *tic20* mutations were identified. For all such double mutants involving *tic20-II* and *tic20-V*, albino plants were observed in the subsequent generation, following self-pollination, at the expected frequency (Table 3.5); for each of the four relevant crosses, ten albino plants were genotyped and shown to be double homozygotes (data not shown). Moreover, these double-homozygous albino plants were not phenotypically different from the corresponding *tic20-I* single-mutant parent (Figure 3.9a). This suggests that atTic20-I shares little or no functional redundancy with the Group 2 proteins.

By contrast, in the F₃ progeny of the *tic20-I tic20-IV* double mutants, no albinos were observed (Table 3.5). Our failure to identify any albinos amongst ~2000 F₃ individuals suggested that the double homozygotes are not viable. To ensure that the double homozygotes were not amongst the green individuals, ~30-60 such plants from each cross were subjected to PCR genotyping; none of them was found to be double homozygous (data not shown). To investigate the possibility that the double homozygous genotypes were embryo lethal, the siliques of individuals that were heterozygous for *tic20-I* and homozygous for *tic20-IV* were investigated. In each case, we observed a significant number of very small, brown aborted seeds, and a much larger proportion of even smaller, white failed ovules (Figure 3.11a). Detailed scoring revealed the frequencies of these defective structures to be ~5-10% and ~40-50%, respectively (Figure 3.11b). These data implied a strong but incomplete defect in female gametophyte development, and furthermore indicated that any double homozygotes that do form arrest during embryo and seed development.

To test the hypothesis that the *tic20-I tic20-IV* double mutations affect female gametophytic transmission, we conducted reciprocal crossing experiments between the double mutants and wild type (Figure 3.11c). The results revealed essentially normal transmission through the male gametes (transmission efficiency was 93.6%), and a strong defect in transmission through the female gametes (34.4%) (Howden *et al.*,

1998). These results are consistent with phenotypic observations made upon analysing the double-mutant siliques (Figure 3.11a), and with the conclusions drawn there from.

Parental Genotype	Parental Generation					
	Selection	Green	Bleached	Green : Bleached	Resistant : Sensitive	
<i>tic20-I-tic20-IV-1</i>	F3	None	473	0	-	-
<i>tic20-I-1 tic20-IV-2</i>	F3	None	748	0	-	-
<i>tic20-I-2 tic20-IV-1</i>	F3	None	857	0	-	-
<i>tic20-I-2 tic20-IV-2</i>	F2	None	664	0	-	-
<i>tic20-I-1 tic20-II-1</i>	F2	None	1244	210	5.91 : 1.0	-
<i>tic20-I-2 tic20-II-2</i>	F2	None	977	234	4.17 : 1:0	-
<i>tic20-I-1 tic20-V-1</i>	F3	None	671	92	7.29 : 1.0	-
<i>tic20-I-2 tic20-V-2</i>	F2	None	681	177	3.85 : 1.0	-

Table 3.5. Genetics analyses of the *tic20* (*I*, *IV*, *II* and *V*) double mutants. Segregation of the albino phenotype in various *tic20-I* double-mutant populations.

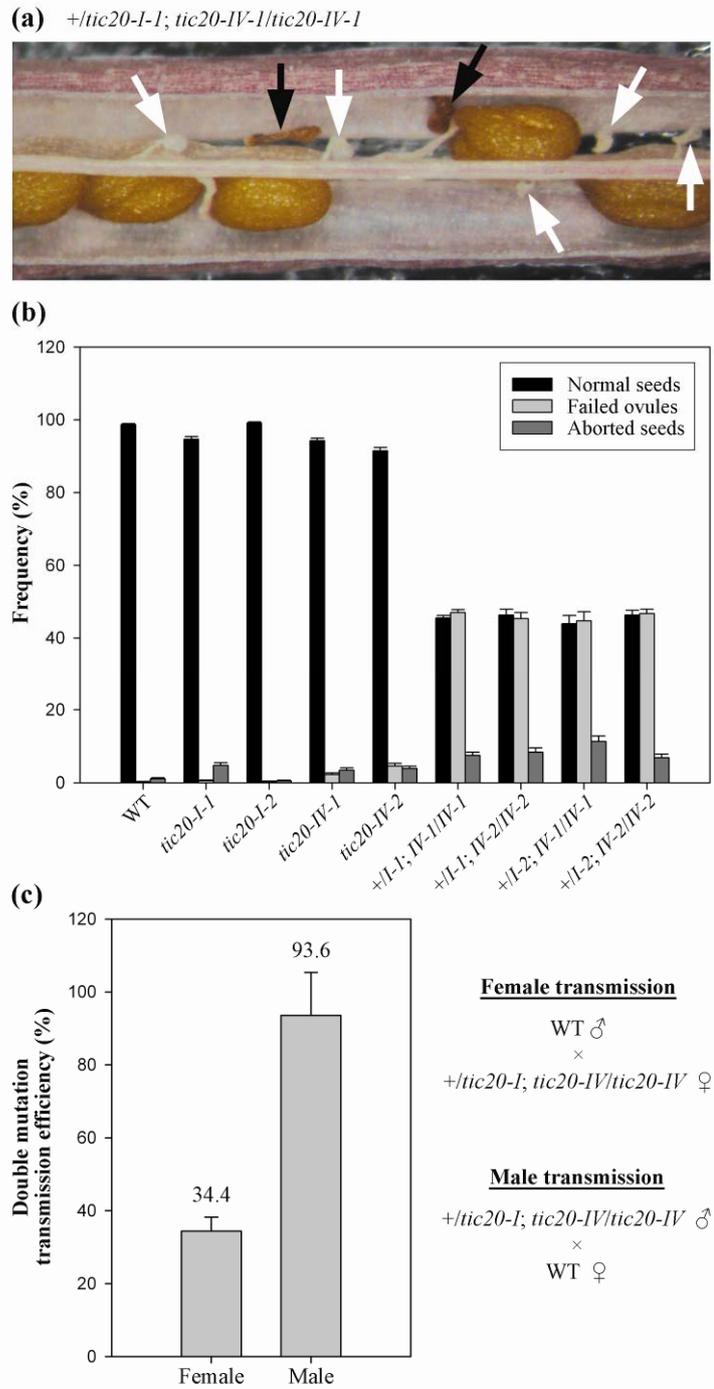


Figure 3.11. Genetic analysis of *tic20-I tic20-IV* double mutants.

(a,b) The siliques of mature, soil-grown plants of the indicated genotypes were opened and inspected. In the siliques of *tic20-I tic20-IV* heterozygous double-mutant plants (genotype: *+/tic20-I; tic20-IV/tic20-IV*), a large proportion of failed ovules and a smaller but significant number aborted seeds was observed. (a) A representative *tic20-I-1 tic20-IV-1* silique is shown. White arrows indicate failed ovules; black arrows indicate aborted seeds.

(b) The frequencies of failed ovules, aborted seeds and normal seeds within the siliques of the double-mutant plants, and in those of the indicated control genotypes, was analysed. The data shown are means (\pm SE) derived from analyses of, on average, 30 siliques (a minimum of 400 seeds) per genotype.

(c) Reciprocal crossing analysis. To assess the efficiency of transmission of the *tic20-I tic20-IV* double mutations, through the female and male gametes, each one of the double-mutant genotypes shown in panel (b) was crossed to wild type, multiple times, in both directions as indicated; simultaneous inheritance of both mutations in the F₁ progeny was assessed by PCR genotyping or by monitoring antibiotic resistance. Transmission efficiencies were calculated for each one of the double mutant genotypes, and then these three values were used to calculate the mean values shown (\pm SE).

3.3.5.5. Chlorophyll Analysis of Single, Double and Triple Mutants

The aim of this experiment was to assess the chlorophyll content of the single, double and triple homozygous mutant of *atTIC20-IV*, *atTIC20-II* and *atTIC20-V*. None of the *tic20-IV*, *tic20-II* and *tic20-V* single mutants was detectably different from wild type. To investigate the possibility that this might be due to functional redundancy, we proceeded to identify a series of double- and triple-mutant combinations. For all tested combinations, doubly or triply homozygous mutant plants could be identified. However, none of these plants exhibited an obvious mutant phenotype (Figures 3.12a). In each case, chlorophyll content in leaves was similar to that in wild type (Figure 3.12b). Moreover, when photosynthetic parameters were assessed by measuring chlorophyll fluorescence, no significant differences from wild-type plants could be detected (Figure 3.12c). These results suggest that the three genes do not play a major role in chloroplast development under the conditions tested. They also imply that they share little or no functional redundancy. It is not surprising that no genetic interactions were detected between *tic20-IV* and the Group 2 mutations, as the *tic20-I* mutations also did not exhibit any detectable interactions with either *tic20-II* or *tic20-V* (Figure 3.9b; Table 3.5). It is possible that our failure to identify a complete knockout for *atTic20-II* is responsible for the lack of any detectable interactions amongst the Group 2 mutants. However, this seems unlikely as in the *tic20-II-1* allele at least, a very strong reduction in transcript abundance was observed (Figure 3.8c). A more likely explanation is that the Group 2 proteins only play a significant role under non-standard conditions that were not assessed in our studies.

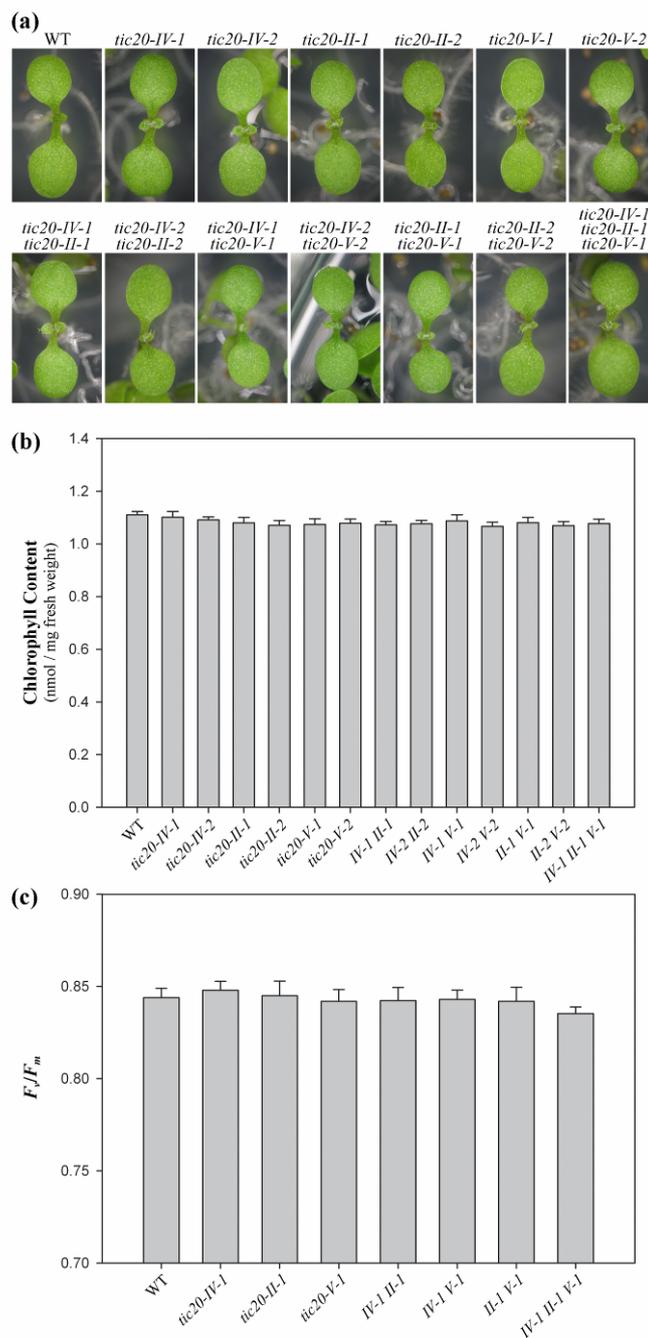


Figure 3.12. Phenotypic analysis of various *tic20* single, double and triple mutant plants.

(a) Homozygous plants of the indicated genotypes were grown side-by-side, *in vitro*, for seven days, and then photographed. Representative plants are shown.

(b) Chlorophyll concentrations in leaves of 30-day-old plants of the indicated genotypes were measured. The plants were grown *in vitro* for 14 days, and thereafter for an additional 16 days on soil. Values shown are means (\pm SE) derived from measurements of 12-15 different plants. Units are nmol chlorophyll *a* + *b* per mg fresh weight.

(c) Analysis of photosynthesis in single- and double-mutant plants. Chlorophyll fluorescence measurements were used to estimate the photochemical efficiency of photosystem II (F_v/F_m). Measurements were done on fully-grown leaves from ten different 30-day-old plants per genotype; all plants were grown under identical conditions, as described in (b). Values shown are means (\pm SE).

3.3.5.6. Complementation Analyses

The aim of this experiment was to further investigate the possibility of redundancy between the different *Arabidopsis* Tic20 homologues. We attempted to complement the *tic20-I-1* mutant line with different Tic20 overexpressor constructs (Table 2.4 cDNA used in this experiment). Heterozygous *+/tic20-I-1* plants were transformed with constructs comprising the *atTIC20-I*, *atTIC20-IV*, *atTIC20-II* and *atTIC20-V* coding sequences under the control of the cauliflower mosaic virus 35S promoter (Appendix 3.4). For each construct, a total of ~8-10 transformants were identified, and from these four independent lines were selected on the basis of overexpression of the relevant *TIC20* transgene, as estimated by RT-PCR. The genotypes of the transformants were all verified by genomic PCR.

As expected, the 35S:*TIC20-I* control construct was very effective at mediating complementation of *tic20-I-1*. This was assessed visually, by making chlorophyll concentration measurements, and by assessing photosynthetic performance through chlorophyll fluorescence measurements (Figure 3.14b). In accordance with the phylogenetic and double-mutant analyses indicating a close relationship between *atTic20-I* and *atTic20-IV* (Figures 3.2 and 3.11), the 35S:*TIC20-IV* construct also mediated significant levels of complementation (Figure 3.13). However, while the 35S:*TIC20-IV* transformants grew to a significantly larger size than the untransformed control plants, this construct was not able to restore normal greening in the *tic20-I* mutant, unlike the control 35S:*TIC20-I* construct. These results indicate significant but incomplete functional redundancy between *atTic20-I* and *atTic20-IV*.

By contrast, the 35S:*TIC20-II* and 35S:*TIC20-V* constructs did not mediate detectable complementation of the *tic20-I-1* albino phenotype (Figure 3.14d), in any of the transformants identified. One caveat is that we were unable to identify transformants (for either of these constructs) that exhibited very high levels of overexpression, in contrast with the situation for the 35S:*TIC20-I* and 35S:*TIC20-IV* constructs discussed above. The reason for this is uncertain, although it may relate to transcript stability. Nonetheless, our genetic analyses (Figures 3.9b and 3.12) do suggest that, even with higher levels of overexpression, a positive complementation result would be unlikely in these experiments.

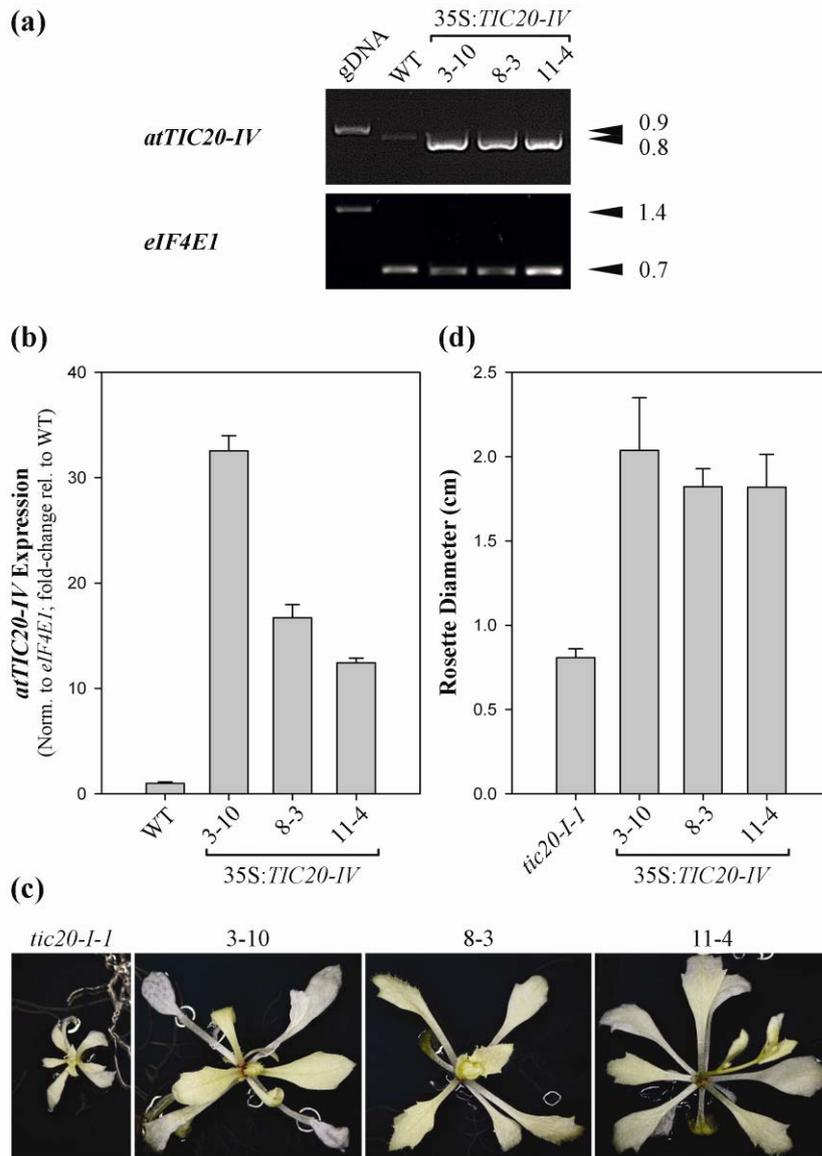


Figure 3.13. Complementation of the *tic20-I-1* mutant by the overexpression of *atTIC20-IV*.

(a,b) Analysis of the expression of the 35S:*TIC20-IV* transgene in three independent transformants (lines 3-10, 8-3 and 11-4). Total-RNA samples were extracted from pools of 10-day-old, T₄ generation seedlings grown *in vitro* that were homozygous for the relevant transgene, as well as from similar wild-type plants; each sample was representative of ~20-30 whole seedlings. Samples were analysed by RT-PCR using the *atTIC20-IV* primers indicated in Figure 5a, using a total of 25 cycles of amplification. Analysis of the housekeeping gene, *eIF4E1*, was used to normalize sample loading. Sizes of the amplicons are indicated to the right of the representative gel image (in kb). The chart shows mean fold-change values (\pm SE), relative to wild type, derived from four independent amplifications (b); the values have been normalized using similarly-derived *eIF4E1* data.

(c) Appearance of the transgenic lines. Populations segregating for the *tic20-I-1* mutation (and which either lacked the 35S:*TIC20-IV* transgene [*tic20-I-1* control], or were homozygous for the indicated transgene [T₄ generation]) were plated on standard MS medium. After five days of growth, homozygous albino individuals were rescued to MS medium containing 3% sucrose and grown side-by-side under standard conditions until they were 38 days old in total. Multiple plants per genotype were then photographed, and the representative images shown were selected.

(d) Rosette size in the transgenic lines. The diameter of the rosette (at its widest point) of each of the plants described in (c), as well as of a number of similar plants, was measured. The values shown are means (\pm SE) derived from ~8-24 different plants per genotype.

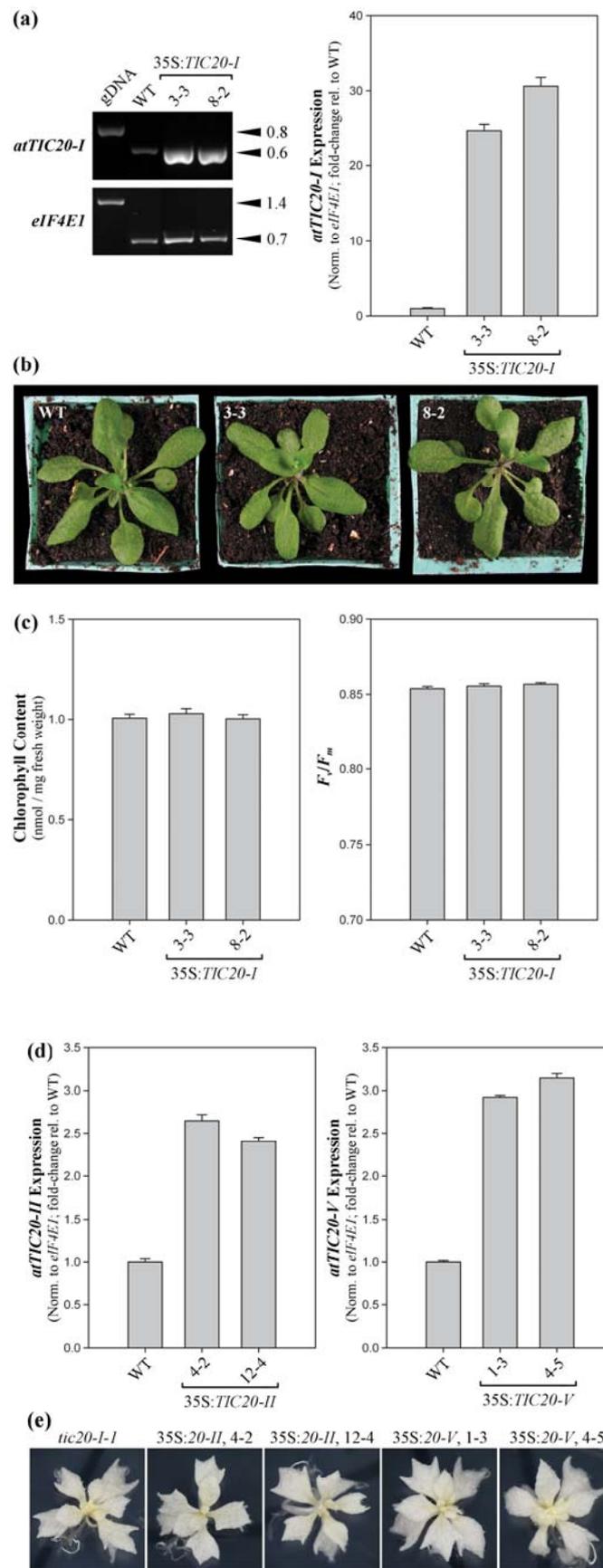


Figure 3.14. Transgenic overexpression of *atTIC20-I*, *atTIC20-II* and *atTIC20-V* in *tic20-I-1* mutant plants. (a,b,c) Analysis of two independent 35S:*TIC20-I* transformants (3-3 and 8-2). (a) Assessment of

transgene expression. Total-RNA samples were extracted from pools of 10-day-old, T₄ generation seedlings grown *in vitro* that were homozygous for the relevant transgene, as well as from similar wild-type plants; each sample was representative of ~20-30 whole seedlings. Samples were analysed by RT-PCR using the *atTIC20* primers indicated in Figure 3.8a, using a total of 25 cycles of amplification. Analysis of the housekeeping gene, *eIF4E1*, was used to normalize sample loading. Sizes of the amplicons are indicated to the right of the representative gel images (in kb). The chart shows mean fold-change values (\pm SE), relative to wild type, derived from four independent amplifications; the values have been normalized using similarly-derived *eIF4E1* data. (b) Appearance of the 35S:*TIC20-I* transgenic plants. Populations homozygous for the *tic20-I-1* mutation and the 35S:*TIC20-I* transgene were plated on standard MS medium, transferred to soil after 14 days, and thereafter grown under standard conditions until they were 28 days old. Transgenics were grown alongside wild-type control plants. Multiple plants per genotype were photographed prior to the selection of the representative images are shown. (c) Chlorophyll content and photosynthetic efficiency in the 35S:*TIC20-I* lines. Wild-type and transgenic *tic20-I-1* mutant plants were grown exactly as described in (b); measurements were taken after 28 days growth. Chlorophyll content values shown are means (\pm SE) derived from measurements of ten different plants (units are nmol chlorophyll *a* + *b* per mg fresh weight). Values for the maximum quantum efficiency of PSII (F_v/F_m) are means (\pm SE) derived from measurements of xx-xx different plants. (d,e) Analysis of two independent transformants for each of 35S:*TIC20-II* (4-2 and 12-4) and 35S:*TIC20-V* (1-3 and 4-5). (d) Assessment of transgene expression. This was conducted exactly as described above in relation to the 35S:*TIC20-I* lines in panel (a). (e) Appearance of the 35S:*TIC20-II* and 35S:*TIC20-V* transgenic plants. Populations segregating for the *tic20-I-1* mutation and which were homozygous for the 35S:*TIC20* transgenes were plated on standard MS medium. Homozygous albino mutants were rescued to MS medium containing 3% sucrose after five days, and were thereafter grown under standard conditions until they were 38 days old. Transgenics were grown alongside *tic20-I-1* control plants. Multiple plants per genotype were photographed prior to the selection of the representative images are shown.

Chapter 4

Molecular Genetic Study of Tic22 Homologues in *Arabidopsis thaliana*

4.1. Abstract

In this study, like the previous Chapter, *Arabidopsis thaliana* has been used as a plant model. Here, I use it to investigate the involvement of Tic22 (translocon at the inner envelope membrane of chloroplasts, 22 kD) in protein import. In *Arabidopsis*, there are two Tic22 homologues termed atTic22-IV and atTic22-III with strong sequence similarity to the pea protein (psTic22). Previously, the Tic22 protein was identified as a candidate of the general protein import apparatus. Tic22 is a peripheral protein of the intermembrane space between the outer and inner envelope membranes and is associated with the outer face of the inner membrane. The topology of Tic22 and its location led to the hypothesis that this protein acts as a link between the TOC and the TIC complexes. It has been proposed that Tic22 and other Tic proteins associate with import components of the outer envelope membrane to form TOC-TIC supercomplexes. Phylogenetic study of Tic22 revealed that these proteins are true homologues of psTic22 protein. The results showed that the Tic22 proteins belong to phylogenetically distinct clade, separate from algae and cyanobacteria proteins. In this study I have used TargetP analysis to predict the subcellular localization of the two *Arabidopsis* Tic22 protein, to the chloroplast. To confirm the TargetP predictions, YFP labelled atTic22 proteins expressed in *Arabidopsis thaliana* protoplasts were analysed. This study allowed us to verify that these proteins are chloroplast proteins. In parallel to YFP analysis, the protein import study was conducted to confirm that these proteins are chloroplast proteins, as the YFP signals for this study were not as good as other proteins analysed (Tic20 homologues for example). To determine *atTIC22* gene expression profiles, publicly-available microarray data were analysed using the Genevestigator tool which revealed expression at different plant developmental stages. Subsequently, real time quantitative RT-PCR revealed that both atTic22 homologue genes are expressed throughout developmental in different tissues. The analysis revealed that; *atTIC22-IV* expression level high (~5-folds) compared to *atTIC22-III*. To determine the functional significance of the gene *in vivo*, T-DNA knockout mutants of Tic22 homologous genes in *Arabidopsis* were identified. Segregation analysis was performed in order ensure the identification of only single-locus insertion lines. In addition, RT-PCR (semi-quantitative) analysis suggested that each of these lines had null mRNA expression. To test the hypothesis that the atTic22 knockouts are important for protein import into chloroplasts, the mutants (and double mutant combinations) were analysed in detail.

Since the single mutants of these genes do not have any visible phenotype, the double mutants for these genes were studied. Phenotypically, double mutants have a pale (yellowish colour) phenotype at early stages of plant development caused. this indicates that there is redundancy between atTic22-IV and atTic22-III. The double-mutant plants were chlorophyll deficient up until the tenth day after germination, but thereafter contained natural pigment level. Electron microscopy revealed that chloroplasts in double mutants were small and under-developed.

Division of labour. The contributors to this Chapter were as follows: Ali Reza Kasmati, Ramesh Patel, Sybille Kubis, Jocelyn Bédard, Morvarid Shirmohammadi and Paul Jarvis. Ramesh Patel assisted me with the identification and characterization of the T-DNA insertion lines. Sybille Kubis assisted me to generate the atTic22-IV and atTic22-III YFP fusion plasmids, and with the transfection and fluorescence microscopy studies. Jocelyn Bédard helped me with cDNA and T-DNA sequences. Morvarid Shir Mohammadi assisted me with identification of double mutants. This research was conducted under the supervision of Paul Jarvis.

4.2. Introduction

In this study, like the previous Chapter, *Arabidopsis thaliana* has been used as a plant model. Here I investigate the involvement of Tic22 (translocon at the inner envelope membrane of chloroplasts, 22 kD) in protein import and chloroplast biogenesis. In the *Arabidopsis* genome two homologous, Tic22 related sequences have been reported to exist, and these are named according to their chromosomal locations (*atTIC22-IV* and *atTIC22-III*). A sequence alignment analysis revealed that atTic22 contains conserved features shared by the psTic22 homologue. These two sequences and psTic22 were reported to be 68.7% and 39.0%, respectively, across the aligning regions. Thus, while it seems likely that atTic22-IV is the direct functional orthologue of psTic22, roles in protein import for the atTic22-III cannot be excluded. Similar to the psTic22 protein, both atTic22 proteins do not possess predicted transmembrane domains. As it was stated before, our knowledge about the TIC complex is very limited, and this fact also to the Tic22 protein. This protein does not possess any specific sequence motifs (for instance, nucleotide binding domains) neither has it any sequences similarity to other components of protein transport system. Tic22 has been suggested to act as a receptor for precursors when they approach from the Toc complex (Kouranov and Schnell, 1997).

Prior study determined the primary structure of Tic22 and its localization within the inner chloroplast envelope (Kouranov and Schnell, 1997). They identified Tic22 as a component of protein import machinery, which can be covalently cross-linked to nuclear-encoded preproteins undergoing import across the envelope (Kouranov and Schnell, 1997). During translocation, it has shown by cross-linking and immunoprecipitation that Tic22 interacts with the components of the TOC translocon of the chloroplast outer envelope. Tic22 cross-links to other TIC and TOC components to form an active super-complex in the chloroplast envelope membrane (Kouranov *et al.*, 1998). It is believed that Tic22, together with other intermembrane space proteins such as Toc64, Toc12 and Hsp70, forms an intermembrane space portion of the translocation complex (Becker *et al.*, 2004b). It is believed that Tic22 might have the binding site for the precursor when they approach from the TOC complex (Kouranov and Schnell, 1997; Kouranov *et al.*, 1999).

The psTic22 protein is peripherally associated with the outer surface of the IEM and to a minor extent to the inner surface of the OEM (Kouranov *et al.*, 1998; Kouranov *et al.*, 1999). It was reported that Tic22 interacts with preproteins before other TIC

complex components, as they emerge from the TOC channel (Kouranov and Schnell, 1997). The localization of Tic22, together with its associations with other Toc/Tic components (Kouranov *et al.*, 1999; Hörmann *et al.*, 2004), suggests that Tic22 acts as a bridge protein between the Toc and Tic translocons, which directing incoming protein, from the outer to the inner membrane (Kouranov *et al.*, 1999; Soll and Schleiff, 2004).

Prior study identified a protein encoded by the gene *slr0924* (Kaneko *et al.*, 1996) in *Synechocystis* PCC6803 (a cyanobacterium) that has significant similarities to the Tic22 subunit (Kouranov *et al.*, 1998). psTic22 contains 19% sequence identity over 176 amino acid residues with the *Synechocystis* (Slr0924) (Reumann and Keegstra, 1999). Recently, *Plasmodium falciparum* Tic22 (pfTic22) was characterized and it was revealed that this protein is associated with apicoplast membranes which are similar to the chloroplast inner envelope membrane. The pfTic22 protein shares 21% identity (40% similarity) with the psTic22 orthologue (Kalanon *et al.*, 2009). Similar to other Tic22 homologues (Kouranov *et al.*, 1998), pfTic22 also has no predicted transmembrane domains. In comparison to Tic22 in plant, the pfTic22 study proposed that this protein perhaps is associated to the apicoplast inner membrane protein, Tic20 (Kalanon *et al.*, 2009).

Deletion mutants and chimeric protein studies revealed that the presequence of Tic22 is required for targeting to the IMS (Kouranov *et al.*, 1999). The protein import of psTic22 was found to be stimulated by ATP and involve the presence of protease-sensitive components on the chloroplast surface. Import studies using an excess of precursor of the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (pSSU) revealed that Tic22 targeting to the intermembrane space does not engage the general protein import pathway used by stromal preproteins (Kouranov *et al.*, 1999). This confirmed that the psTic22 presequence does not operate as a stromal transit peptide and that psTic22 is targeted to the chloroplasts intermembrane space by a novel import pathway which is different from known pathways (Kouranov *et al.*, 1999).

4.3. Results

4.3.1. Bioinformatics

4.3.1.1. *Arabidopsis thaliana* Tic22 Homologous Protein Alignments

The aim of this experiment was to reveal the similarity between two proteins in *Arabidopsis* related to psTic22. According to our database searches, in *Arabidopsis* there are two homologues (atTic22-IV and atTic22-III) which may function in a similar way to psTic22 in protein import pathways. Full-length, sequenced cDNA clones are available for each gene (accession numbers AK118805 and NP-189013, respectively), confirming that they are expressed in plants. Protein sequences predicted using these cDNA sequences were analysed by *in silico* experiments. Analysis of sequence alignments made using ClustalW revealed high sequence similarity of psTic22 to atTic22-IV (~68.7% identity), whereas atTic22-III shares much less similarity to psTic22 (39.0%), as shown in Figure 4.1a. The atTic22 homologues have been compared to other species such as the rice homologues (*Oryza sativa*) proteins, osTic22 (BAD-35192) and osTic22 (XP-477538) (Figure 4.1b). I repeated the alignments described above using the BioEdit program shown in Figure 4.2 as the MegAlign program does not produce alignment in format that easily presentable, to reveal that the *Arabidopsis* Tic22 homologues and pea do share conserved sequences.

(a)

Percent Identity

1	2	3	
	68.7	39.0	1
		33.8	2
			3

psTic22 protein
atTic22-IV protein (AK-118805)
atTic22-III protein (NP-189013)

(b)

Percent Identity

1	2	3	4	
	38.7	56.6	32.9	1
		36.4	51.8	2
			32.7	3
				4

osTic22 protein (NP-001059394)
osTic22 protein (BAD-35192)
atTic22-IV protein (AK-118805)
atTic22-III protein (NP-189013)

Figure 4.1. Comparison of the psTic22 and atTic22 protein sequences.

(a) Percent amino acid identities shared between psTic22 and the *Arabidopsis* Tic22 homologues. Values were calculated using MegAlign (ClustalW).

(b) Comparison of the atTic22 and osTic22 protein sequences. Percent amino acid identities shared between the *Arabidopsis* and rice Tic22 homologues. Values were calculated using MegAlign (ClustalW).

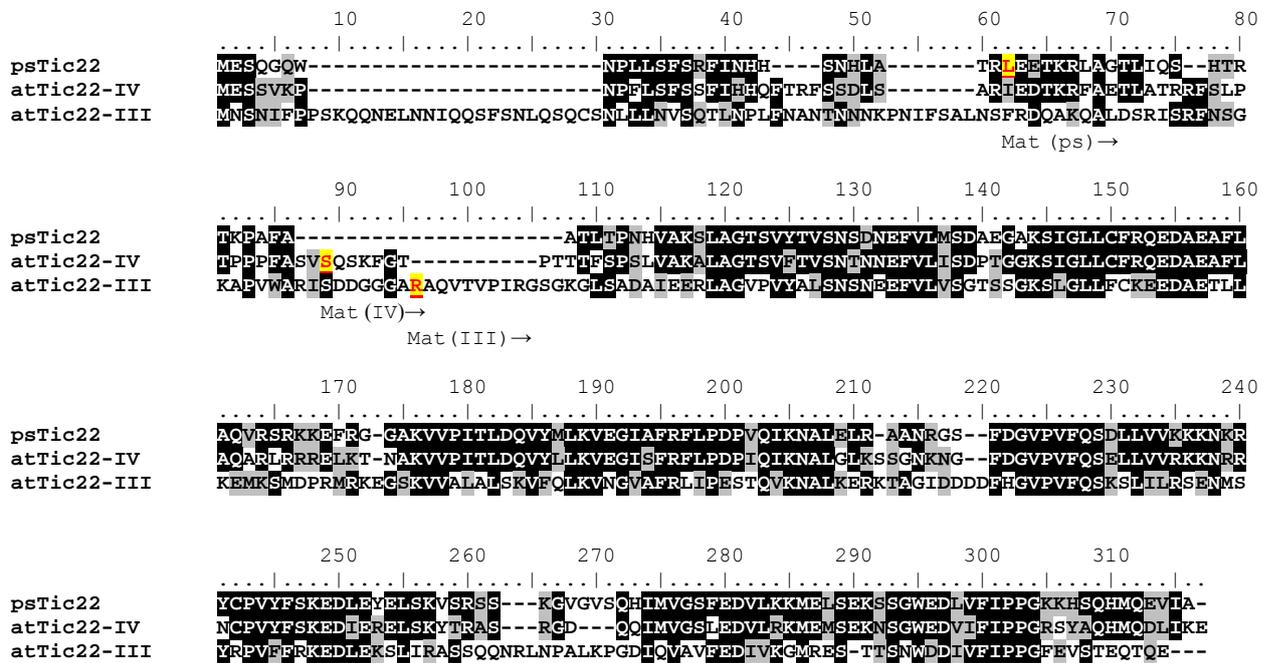


Figure 4.2. Amino acid alignment of psTic22 with the two *Arabidopsis* homologues, Tic22-IV and atTic22-III. The sequences were aligned by Clustal W using the BioEdit program. Mat→ represents the starting residue of the mature form of each protein, as predicted by TargetP. Additionally, the first residue of each predicted mature sequence is underlined (29, 60, and 90 for psTic22 and atTic22 homologues, respectively).

4.3.1.2. Phylogenetic Analysis and Comparison with Other Organisms

This aim of this experiment was to compare atTic22 homologues to other species (Table 4.1) and to assess whether atTic22 homologues are truly homologous to psTic22 (Figure 4.3). The amino acid sequences of Tic22 homologues and from many species (Table 4.1) were retrieved by BLAST searching. First, we manually aligned a representative set of Tic22 and other related proteins to the prokaryotic sequence, Figure 4.3.

In comparison to the atTic22-IV proteins, the atTic22-III is more divergent but it is still within the same clade. In parallel to comparison study (Figure 4.1a) the distances between atTic22-III and psTic22 is much larger which indicates that these proteins have less similarity. Additionally, Figure 4.3 shows that the atTic22 family diverged into two distinct sub-clades (atTic22-IV and atTic22-III). This suggests that there are two sub-types of atTic22-related protein in plants.

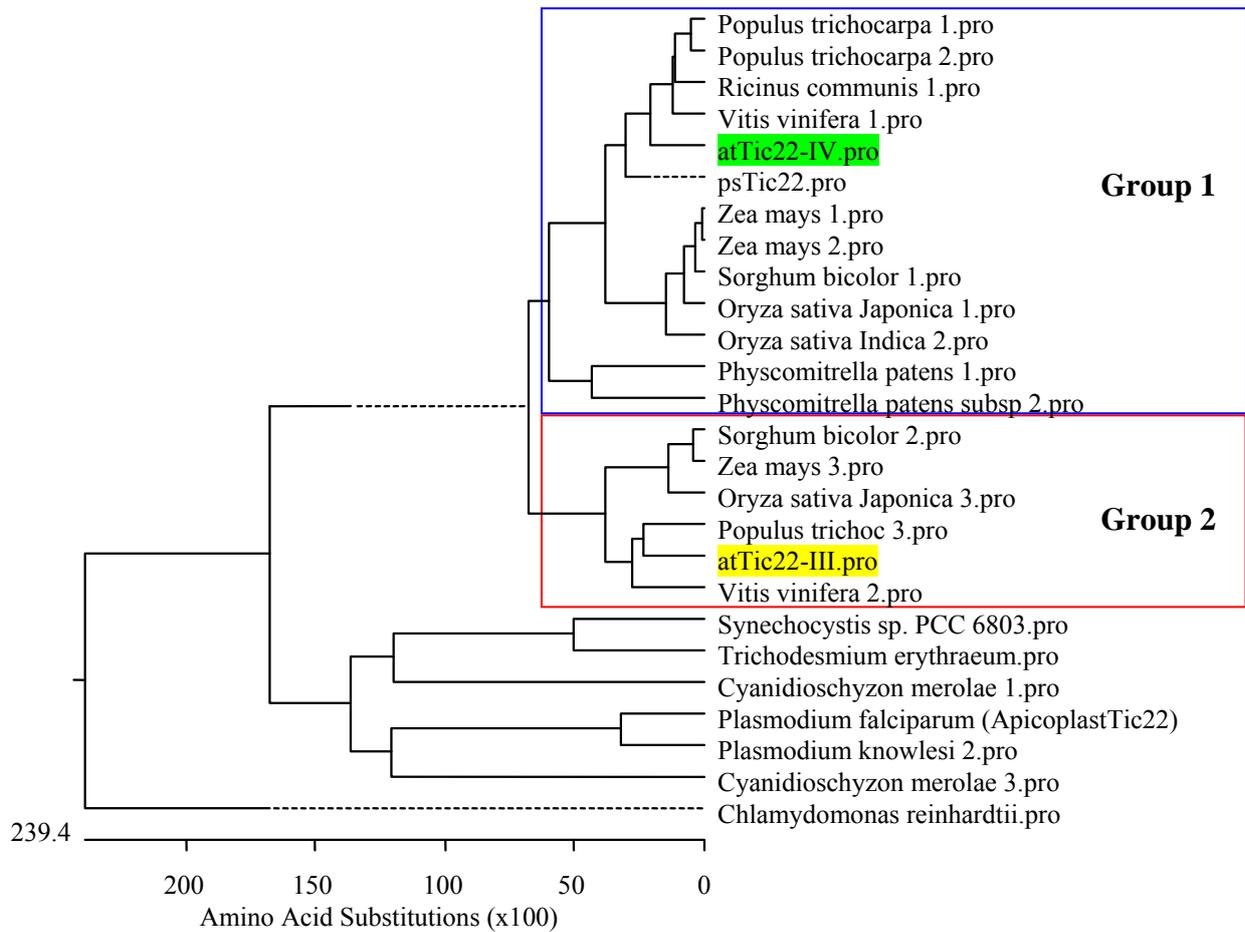


Figure 4.3. Phylogenetic analysis of Tic22-related sequences from different species.

Amino acid sequences of the TargetP-predicted mature regions of the indicated Tic22 homologues were aligned and used to produce a phylogenetic tree. The two major clades are termed Group 1 and Group 2, as indicated. Protein and accession numbers of the sequences used are provided in Table 4.1. Species of origin is indicated with the following prefixes: ps, *Pisum sativum*; at, *Arabidopsis thaliana*; os, *Oryza sativa* (japonica subspecies); osi, *Oryza sativa* (indica subspecies); pp, *Physcomitrella patens*.

Species Name	Accession Number	Number	Name
<i>Arabidopsis thaliana</i>	AK-118805		atTic22-IV
<i>Arabidopsis thaliana</i>	NP-189013		atTic22-III
<i>Chlamydomonas reinhardtii</i> ¹	XP_001692709		
<i>Cyanidioschyzon merolae</i> ²	CMC181C	1	
<i>Cyanidioschyzon merolae</i> ²	CMJ012C	2	
<i>Cyanidioschyzon merolae</i> ²	CMJ105C	3	
<i>Oryza sativa Japonica</i>	BAD35192	1	
<i>Oryza sativa Indica</i>	EEC80218	2	
<i>Oryza sativa Japonica</i>	NP_001059394	3	
<i>Physcomitrella patens</i>	XM_001766060.1	1	
<i>Physcomitrella patens subsp</i>	XP_001780590	2	
<i>Pisum sativum</i>	AAC64606		psTic22
<i>Plasmodium falciparum</i> ³	XP_001351847	1	pfTic22
<i>Plasmodium knowlesi</i>	XP_002259519	2	
<i>Populus trichoc</i>	XP_002326937	1	
<i>Populus trichoc</i>	XP_002302438	2	
<i>Populus trichoc</i>	XP_002301750	3	
<i>Ricinus communis</i>	XP_002510224	1	
<i>Ricinus communis</i>	XP_002518637	2	
<i>Synechocystis sp. PCC 6803</i> ⁴	NC_000911.1	1	
<i>Trichodesmium erythraeum</i> ⁵	Q115A9	2	
<i>Sorghum bicolor</i>	XP_002438041	1	
<i>Sorghum bicolor</i>	XP_002459723	2	
<i>Vitis vinifera</i>	XP_002284687	1	
<i>Vitis vinifera</i>	XP_002268264	2	
<i>Zea mays</i>	ACF85455	1	
<i>Zea mays</i>	NP_001151773	2	
<i>Zea mays</i>	NP_001136600	3	

¹ Green Alga

² Red Algae

³ Apicoplast Tic22

⁴ slr0924 (Cyanobacteria)

⁵ IMS101 (Cyanobacteria)

Table 4.1. The organism abbreviation, the Tic22-related amino acid sequences used for the phylogenetic analysis.

4.3.1.3. Prediction of Protein Localization (by TargetP)

The aim of this experiment was to develop a better understanding of subcellular localization of atTic22 homologues. Subcellular localization of proteins is essential for the structure and function of the cell. Consequently, awareness of the subcellular localization of proteins is crucial to recognize their roles and interacting partners in cellular metabolism (Kumar *et al.*, 2002; Huh *et al.*, 2003). The TargetP program prediction revealed that both *Arabidopsis* Tic22 homologues have a transit peptide with high confidence. The results are shown in Table 4.2 (all raw data for this experiment are attached in appendix 4.1).

Amino acid sequences of psTic22 and the two atTic22 homologues were chosen for the TargetP analyses. Predictions for chloroplast localization were achieved by using this web-site (<http://www.cbs.dtu.dk/services/TargetP/>), and the results are shown in Table 4.2. The TargetP prediction provides a Reliability Class (RC) score according to the difference between the highest and the second-highest network output score; this feature specifies the level of certainty in the prediction for a particular sequence. The lower the RC, the better the prediction which is more reliable (Table 4.2). Prediction by TargetP is not 100% reliable since psTic22 has been shown to be a component of the protein import machinery at the chloroplast inner envelope membrane (Kouranov *et al.*, 1998), whereas here it is predicted to be located in mitochondria (Table 4.1).

Protein	Predicted Location	Reliability Class	TP Length (aa)
psTic22	Mitochondria	5	28
atTic22-IV	Chloroplast	3	59
atTic22-III	Chloroplast	2	96

Table 4.2. TargetP was used to predict the location of each protein; the Reliability Class and the predicted length of the transit peptide (TP) are also shown.

4.3.2. Protein Localization by YFP

The aim of the present work was to provide experimental support for the TargetP predictions. The subcellular localization of the *Arabidopsis* Tic22 homologues was assessed by the analysis of YFP fusion proteins. To this end, full-length coding sequences for each of both *Arabidopsis* Tic22 genes (Table 2.4 cDNA) were inserted into the p2GWY7 vector by Gateway recombination cloning (Karimi *et al.*, 2005); this vector adds a C-terminal YFP tag (Appendix 3.3, Supplementary Figure 3.1). *Arabidopsis* protoplasts were then transfected using these constructs, and analysed by fluorescence microscopy (Figure 4.4). The red fluorescent signal of chlorophyll shows the location of the chloroplasts in each case. For both fusion proteins, the yellow-green fluorescence of YFP was observed in a faint ring-like pattern around the periphery of each chloroplast. However, these results do not strongly support the hypothesis that *Arabidopsis* Tic22 homologues are localized in the chloroplast inner envelope membrane. The patterns observed were not as clear as the distributions seen for atTic110:YFP and atTic40:YFP fusion proteins in a previous, experimentally very similar study (Bédard *et al.*, 2007), perhaps indicating that the YFP tag interferes with targeting to the chloroplast envelope. The lack of clear localization might be due to atTic22 using different protein pathway compared to the atTic20 proteins or the large YFP tag disrupting targeting of atTic22 which is essential for intermembrane space localization.

In comparison to this study (atTic22-YFP), recently a different group attempted to localize the *Plasmodium falciparum* Tic22 homologue using a pTic22-GFP fusion protein to apicoplasts using an Hsp86 promoter (Kalanon *et al.*, 2009). However, they failed to localize this protein to apicoplasts. The authors proposed that the proper localization of the pTic22-GFP protein might be blocked in the protein transport pathway due to the high expression of the Hsp86 promoter, or to the large GFP fusion disturbing the protein-targeting pathway. This indicated that Tic22 localization in different species using GFP or YFP is not successful, perhaps due to the unique properties of the protein's transport pathway.

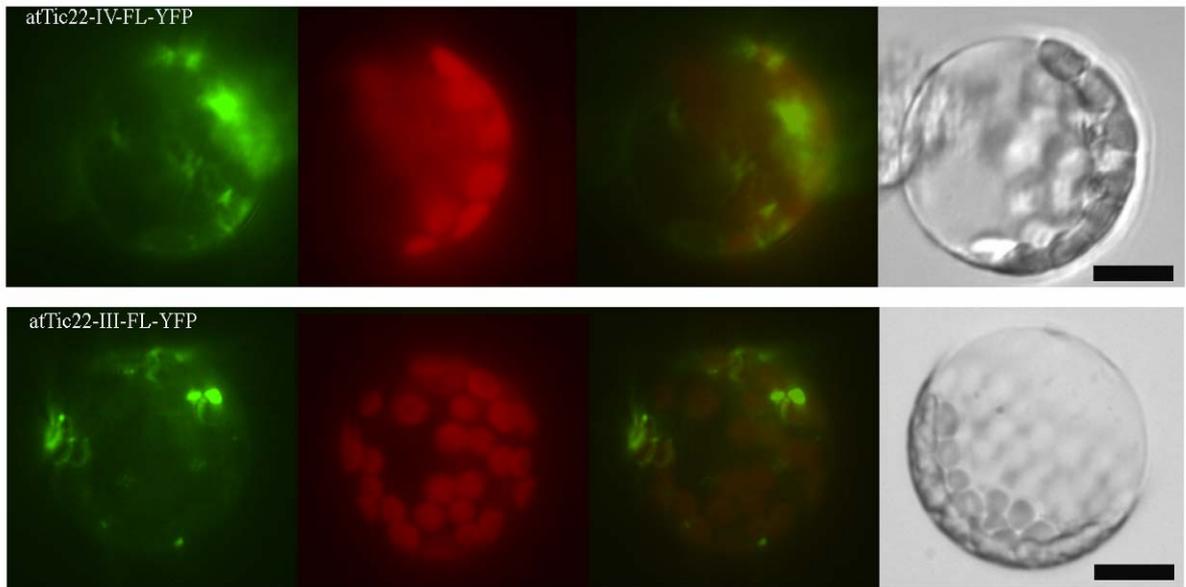


Figure 4.4. Subcellular localization of the *Arabidopsis* Tic22 proteins as assessed by YFP fusion-protein analysis.

Wild-type *Arabidopsis* protoplasts were transfected with the indicated plasmids (atTic22-IV:YFP and atTic22-III:YFP) and then analysed for YFP fluorescence (green, left panels) and chlorophyll autofluorescence (red, centre-left panels), as well as under bright field illumination (right panels). In addition, an overlay of the YFP and chlorophyll images is presented (centre-right panels). In both cases, YFP fluorescence were not as clear as distributions seen for atTic20:YFP fusion proteins in previous Chapter (Chapter 3, Figure 3.4). Scale bars = 10 μ m.

4.3.3. Protein Import

As the atTic22-YFP fusion (both proteins) study was inconclusive, *in vitro* chloroplast protein import of atTic22 experiments have been performed using ³⁵S-labelled precursor proteins. Previously *in vitro* study revealed that psTic22 imports into pea chloroplasts at a very low velocity (Kouranov *et al.*, 1999). In this study I used the precursor of the small subunit of Rubisco (preSSU) as an internal control. This experiment was carried out with post-import thermolysin treatments to confirm the chloroplast location of the atTic22 proteins (Aronsson and Jarvis, 2002; Kubis *et al.*, 2003). As shown in Figure 4.5, the atTic22-IV protein was imported into wild-type chloroplasts in parallel to SSU. In summary, this result showed that both atTic22-IV (and SSU) imported into wild-type chloroplasts. Since there was difficulty with atTic22-III protein import, this protein was eliminated from this study at this stage. In future it will be necessary to attempt to import this protein into wild-type chloroplasts to confirm that this protein is also a chloroplast protein.

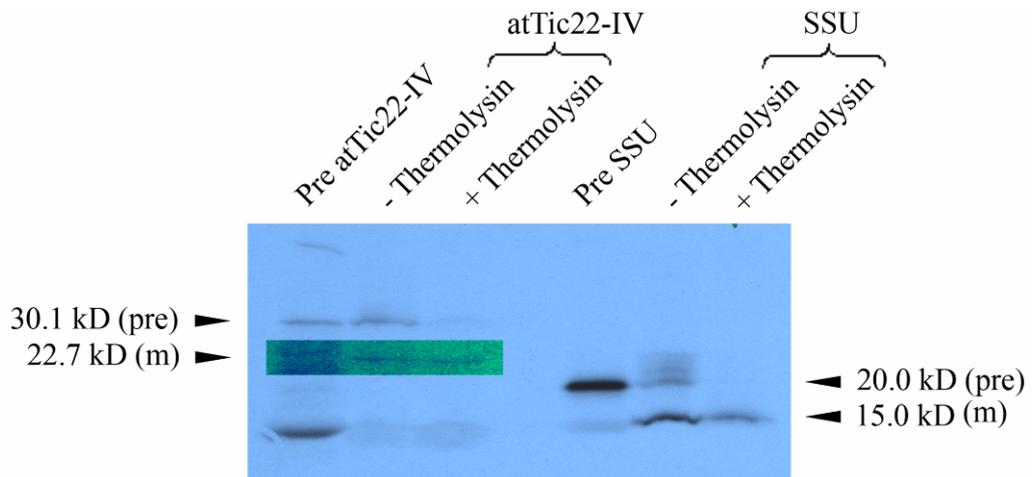


Figure 4.5. Protein import into wild-type chloroplasts.

Isolated wild-type chloroplasts were incubated with *in vitro* translated, [³⁵S] methionine-labelled preTic22-IV and preSSU in the presence of 5 mM ATP for 12 minutes import and 20 minutes thermolysin treatment. Import reactions were analysed by SDS-PAGE, visualized using a phosphor-imager. Representative gel image is shown above. The area around the mature form of atTic22-IV was enhanced for contrast using Photoshop.

4.3.4. Genevestigator Analysis

To obtain information on *atTIC22* gene expression at different plant growth stages, publicly-available microarray data were analysed using the Genevestigator tool (Zimmermann *et al.*, 2004). A developmental time course revealed that both *Arabidopsis TIC22* genes are expressed throughout the plant's life-cycle at very low level compared to control genes (Figure 4.6). Two TOC-related *Arabidopsis* genes, (*atTOC33* and *TOC34*) as well as the gene for a major component of the TIC complex, *atTIC110*, were used as a controls in this experiment (Figure 4.6). As indicated in Figure 4.6, the control genes *atTOC33* and *TIC110* were expressed most strongly, whereas *atTIC22-IV* was expressed at very low levels except in bolting where the expression level of *atTIC22-IV* was similar to *atTOC34*. In comparison to the control genes, the *atTIC22-III* gene exhibited an especially low expression level. However, the expression levels of *atTIC22-IV* in most of developmental stages are similar to a control gene *atTOC34*. These results might indicate that these genes are not especially important in protein import and development in *Arabidopsis*.

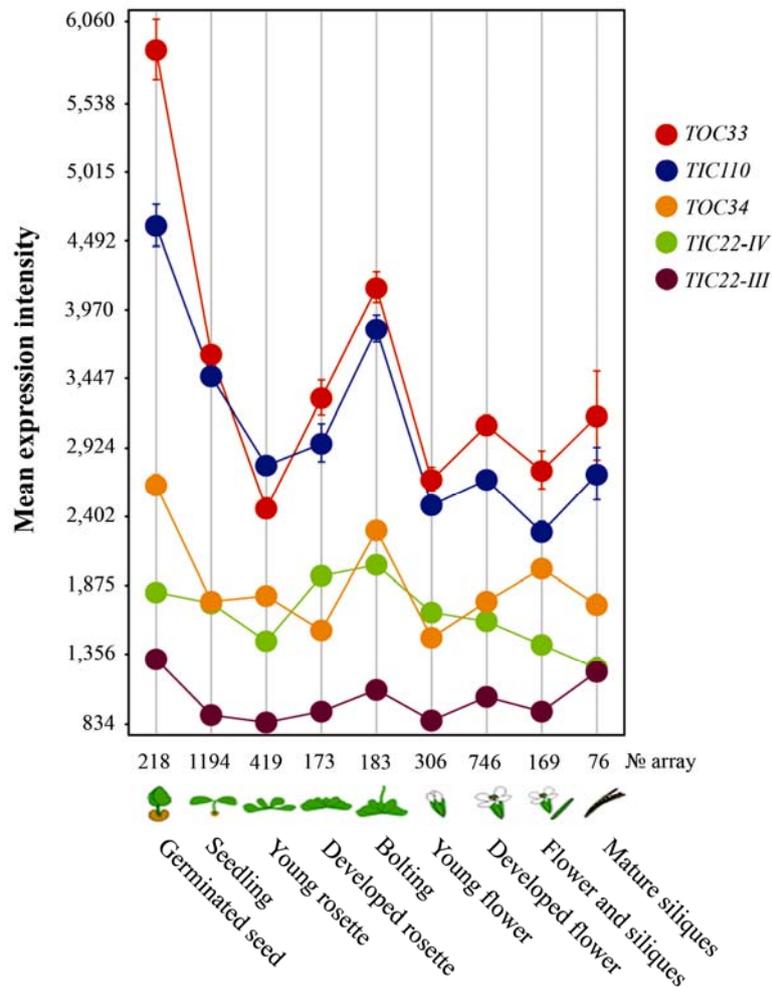


Figure 4.6. Expression of the *Arabidopsis* Tic22 genes using publicly-available microarray data. Affymetrix GeneChip data were analysed and retrieved using the Genevestigator V3 analysis tool (<https://www.genevestigator.ethz.ch/>) (Zimmermann *et al.*, 2004; Grennan, 2006). Presented data were prepared using the Meta-Profile Analysis tool using the Development representation in scatter-plot format. Data from all high-quality ATH1(22k) arrays were analysed; this amounted to a total of 3110 arrays. Values shown are means (\pm SE). The total number of arrays used to derive each data point shown is indicated. Data representations were exported from Genevestigator in Encapsulated PostScript format, and then compiled and annotated using appropriate graphics software. The genes analyzed were as follows: *atTIC22-IV* (At4g33350; green); *atTIC22-III* (At3g23710; purple); *atTOC33* (At1g02280; red); *atTOC34* (At5g05000; orange); and *atTIC110* (At1g06950; blue).

4.3.5. Expression Profiles of *Arabidopsis* TIC22 Homologues

To gain insight into the functional relationships between the *Arabidopsis* TIC22 homologues, their developmental and tissue-specific gene expression patterns were studied by quantitative real time RT-PCR (Figure. 4.7). Samples analysed included rosette leaves, siliques and roots from mature plants. Additionally, I analysed whole 5-day-old seedlings grown in the dark (5dD) or in the light (5dL). The latter sample was compared to 14-day-old, light-grown seedlings (14dL) in order to study a developmental difference.

As expected, the results indicated that *atTIC22-IV* expression was higher than *atTIC22-III* in photosynthetic tissue, and weak in non-photosynthetic growing tissue, such as 5-day-old etiolated plants and roots (Figure 4.7). The overall level of *atTIC22-IV* expression is much higher than *atTIC22-III* expression, at all developmental stages. According to the Genevestigator study, this suggests that *atTIC22-IV*, is relatively more important than *atTIC22-III*. It is possible that *atTIC22-IV* is more important for photosynthetic development than non-photosynthetic development.

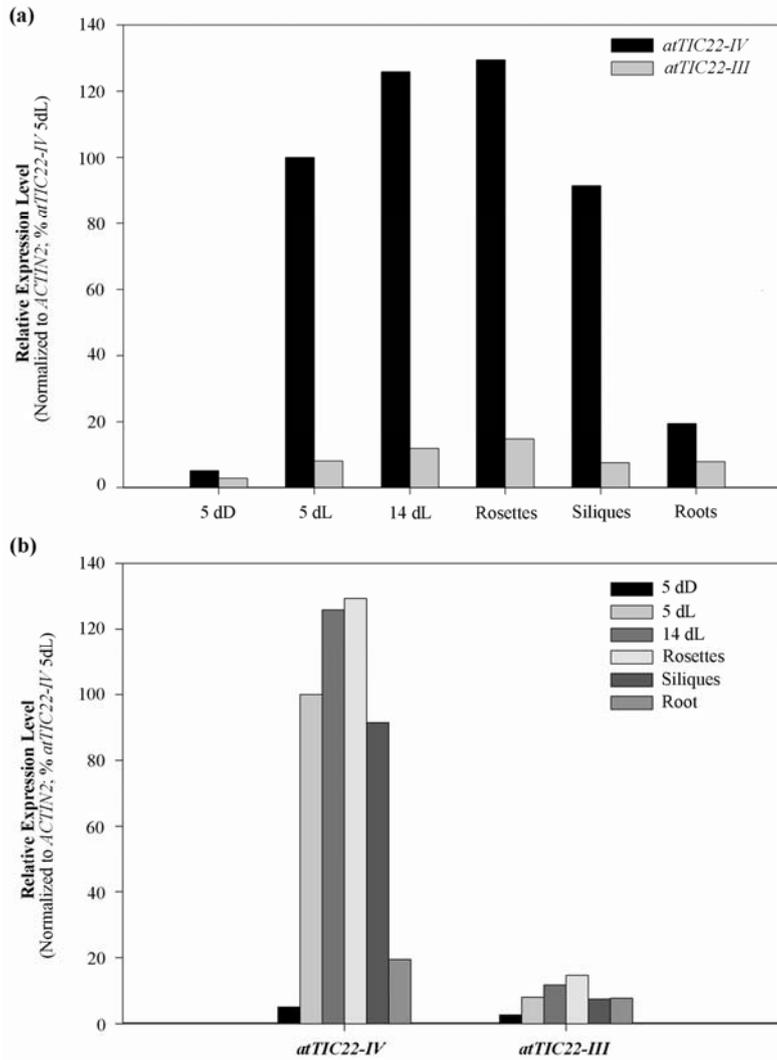


Figure 4.7. Quantitative RT-PCR analysis of the expression of the Tic22 genes in different tissues and at different stages of development.

Total RNA isolated from whole seedlings grown *in vitro* for five days in the dark (5dD), or five and 14 days in the light (5dL and 14dL, respectively), as well as from three different organs of 28-day-old plants grown on soil (roots, rosette leaves and floral buds). The RNA samples were representative of ~10-30 seedlings (5dD, 5dL and 14dL), or 5-25 mature plants (roots, rosettes and buds). Samples were analysed using quantitative real-time RT-PCR. Using the results, ΔC_T values were calculated for both Tic22 gene, relative to the control gene, *ACTIN 2* (At3g18780). These values were then expressed as percentages of the ΔC_T value for the most highly-expressed gene, *atTIC22-IV* (5-day-old in the light). The data shown are means derived from three independent amplifications, and reflect the relative expression levels of the four genes. Panels (a) and (b) contain the same data presented in different ways, to aid interpretation.

4.3.6. Mutant Analysis of *Arabidopsis* TIC22 Genes

The aim of this experiment was to analyse the functional significance of the *Arabidopsis* TIC22 homologues *in vivo*, and to identify two independent T-DNA insertion mutants for each of the two genes. Two independent T-DNA insertions for both genes (*atTIC22*) have been identified. All of the T-DNA insertion sites were confirmed by genomic PCR, and by the sequencing of the T-DNA/gene junctions at one or both sides in each case, as indicated (Figure 4.8).

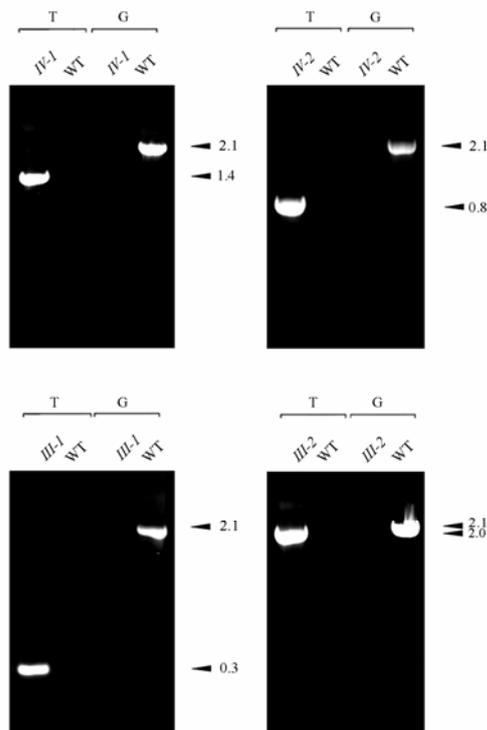


Figure 4.8. Analysis of the Tic22 T-DNA mutants by genomic PCR.

Genomic DNA samples extracted from wild type and putative homozygous mutants were analysed by PCR. Appropriate T-DNA- and *atTIC22*-gene-specific primers were employed. Two different primer combinations were used in each case: the first ('T') comprised one T-DNA border primer and one gene-specific primer (LB + reverse: *tic22-IV-1*, *tic22-IV-2*, *tic22-III-1* and *tic22-III-2*); the second ('G') comprised two gene-specific primers flanking the T-DNA insertion site. The PCR products were resolved by agarose gel electrophoresis, and visualized by staining with SYBR Safe. Amplification using "T" indicated the presence of the mutant allele, whereas amplification using "G" indicated the presence of the wild-type allele; amplification with the former but not the latter demonstrated that the plant was homozygous mutant. The genotype names are shortened as follows: 'IV-1' indicates *tic22-IV-1*; 'IV-2' indicates *tic22-IV-2*; and so on. Sizes of the amplicons are indicated at right (in kb).

4.3.5.1. Segregation Analysis of *Arabidopsis* T-DNA Mutants

Segregation analysis was performed in order to confirm the identification of only single-locus insertion lines (Mendelian ratio as described in Chapter 3), Table 4.2. As is normal for T-DNA transformation, some of these transformed plants contain more than one T-DNA insertion. Further segregation analysis in later generations enabled the identification of homozygous lines for analysis, and the zygosity of these lines was confirmed by genomic PCR (Figure 4.8).

After extensive genetic analysis, genetically “clean” (single-locus insertion lines) lines were identified for most insertions; representative data are shown in Tables 4.2. Plating the *tic22-IV-1* mutant line (Salk-022794) on kanamycin-containing medium revealed that this line has silencing difficulties due to the kanamycin resistance marker. Plating a heterozygous *tic22-IV-2* mutant line (Gabi-Kat_710-E01) on sulfadiazine revealed that all plants were resistant; therefore, it was presumed that this line was not genetically clean. Thus, that plant was back-crossed to wild type in order to clean this line. In the F₂ generation, a 3 resistant : 1 sensitive segregation ratio was observed, and a homozygous line for *tic22-IV-2* (Gabi-Kat_710-E01) was found. The genotype of all lines were also confirmed by PCR and they were homozygous for the T-DNA insertion (Figure 4.8).

Genotype	Generation	Selection	Total resistant		Resistant : Sensitive	χ^2 -value ^c	p-value ^c
			Green	Sensitive			
<i>tic22-IV-1</i> ^a	T ₅	kanamycin	-	-	-	-	-
<i>tic22-IV-2</i> ^b	F ₂	sulfadiazine	61	15	4.07 : 1	1.123	0.289
<i>tic22-III-1</i>	T ₃	sulfadiazine	122	41	2.98 : 1	0.002	0.964
<i>tic22-III-2</i>	T ₃	sulfadiazine	128	42	3.05 : 1	0.008	0.929

Table 4.2. Genetic analyses of the atTic20 homologue mutants. Segregation analysis of the T-DNA-associated antibiotic resistance marker in each one of the Tic22 mutants.

^a All seedlings were classified as green due to scoring difficulties caused by the gene silencing of the kanamycin resistance marker.

^b These plants were derived from a back-cross to Col-0 wild type.

^c Chi-square tests evaluated “goodness of fit” of the observed ratios close to 3 : 1; degrees of freedom = 1.

4.3.5.2. Analysis of T-DNA Insertion Mutants by RT-PCR

To assess the effect of each T-DNA insertion on *TIC22* gene expression, RT-PCR analysis was conducted in each case (Figure 4.9). The results confirmed that the relevant full-length mRNA was absent for all of the mutants and so they are considered to be knockout alleles that produce no mRNA (Figure 4.9). The mRNA samples for RT-PCR used for this experiment were derived from 10-day-old homozygous individuals grown *in vitro*. Amplification was performed using gene specific primers and products were visualized by staining with ethidium bromide following agarose gel electrophoresis. Additionally, the *atTOC33* and the *eIF4E1* genes were used as controls in each experiment. For every reverse transcription (RT) sample. A “water” control (lacking template) was also included for all primer pairs, to verify the absence of contamination in the reagents.

4.3.5.3. The Analysis of *Arabidopsis tic22* Double Mutants

None of the *tic22-IV* and *tic22-III* single mutants was detectably different from wild type. Therefore, double knockout mutants were generated in order to investigate the possibility that this might be due to functional redundancy. In order to analyse functional relationships between *atTIC22-IV* and *atTIC22-III*, homozygotes (both alleles) were crossed to the other mutants. In F2 generations from each cross, the pale phenotypes were observed. To ensure that the phenotypes were genuine, individual plants from each cross were subject to PCR genotyping (data not shown). Twenty five individual plants from each F3 were genotyped by PCR (for both *tic22-IV* and *tic22-III* alleles), and all were found to be homozygote for both alleles. All F3 progeny produced pale phenotypes at 5-days-old; after 5 days the double-mutant phenotypes gradually became like wild-type. The results revealed that the double homozygous genotype caused no severe impact on growth, but indicated that there is indeed functional redundancy between these genes.

4.3.5.4. Chlorophyll Analysis of Single and Double Mutants

The aim of this experiment was to assess the chlorophyll content of the single and double homozygous mutants of *atTIC22-IV* and *atTIC22-III*. The single mutants of these two genes exhibited no obvious visible phenotypes, whereas double homozygous mutants of all combinations displayed pale phenotype at early stages of plant development (Figure. 4.10). The phenotype analyses may suggest that these genes may not have an important role in later stages of chloroplast development. Chlorophyll measurements were used to analyse the phenotypes of the double mutants. Chlorophyll content was assessed by performing a time-course analysis, with measurements on the 5th day, 7th day, 10th and 14th day after germination. The 5-day-old double mutant plants contained reduced chlorophyll concentrations (Figure 4.10). Although only moderate chlorophyll deficiencies were found in these double mutants, this phenotype was clearly apparent in all three double homozygous knockout mutants (Figure 4.12a and b). As this phenotype observed only at early stage of plant development, the plastid ultrastructure was analysed using transmission electron microscopy (Figure 4.11a and b). The cotyledons of 5-day-old (Figure 4.11a and b), *in vitro* grown plants were analysed.

However, as the plants grew older, the chlorophyll content of these double mutants increased, and by day 20 there is no reduction in chlorophyll concentration relative to wild type.

As indicated in Figure 4.12a and b, SPAD readings and HandyPea of chlorophyll contents in leaves were similar to wild-type in all single and double mutants. Individual leaves of ten plants per genotype were analyzed to characterize the phenotype in each case; as Figure 4.12a and b shows there is no significant difference between the genotypes. Taken together, these data may support that the function of Tic22 is not so important during later chloroplast biogenesis and development (Figure 4.12a and b).

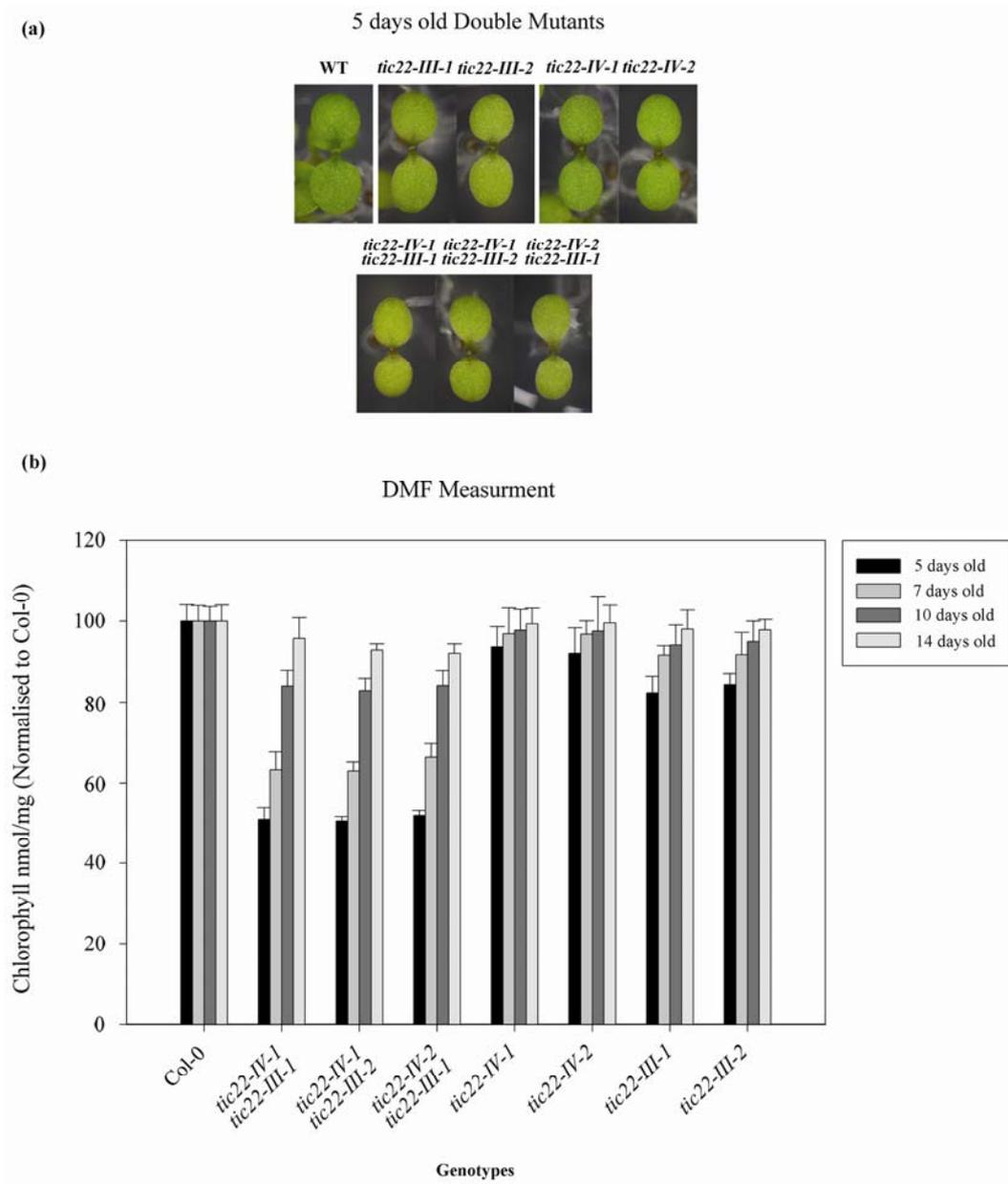
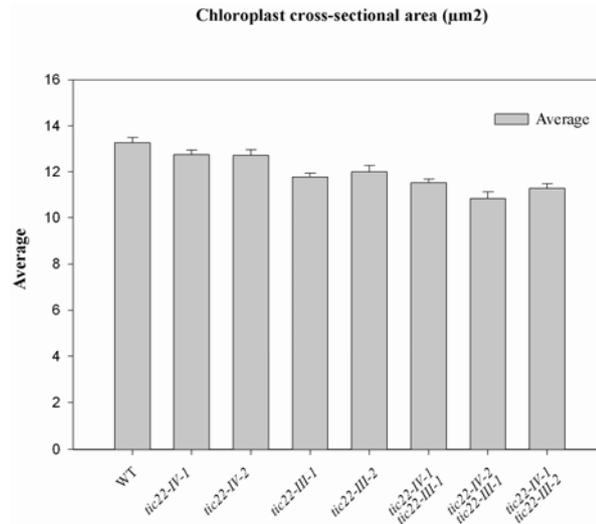


Figure 4.10. Phenotypic analysis of various *tic22* single and double mutant plants.

(a) Homozygous plants of the indicated genotypes were grown side-by-side, *in vitro*, for five days, and then photographed. Representative plants are shown.

(b) Chlorophyll concentrations in single- and double-mutant (homozygous) of 5, 7, 10 and 14-day-old plants (whole plants) of the indicated genotypes were measured. The plants were grown *in vitro* for 14 days under identical conditions (In Percival). Values shown are means (\pm SE) derived from measurements of three replicates. Units are nmol chlorophyll *a* + *b* per mg fresh weight.

(a)



(b)

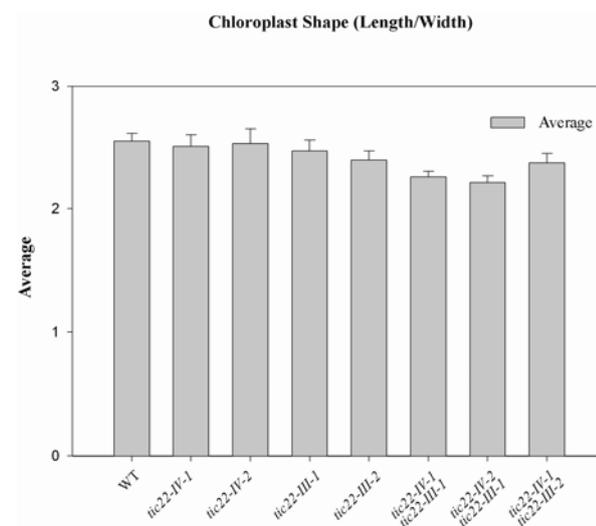
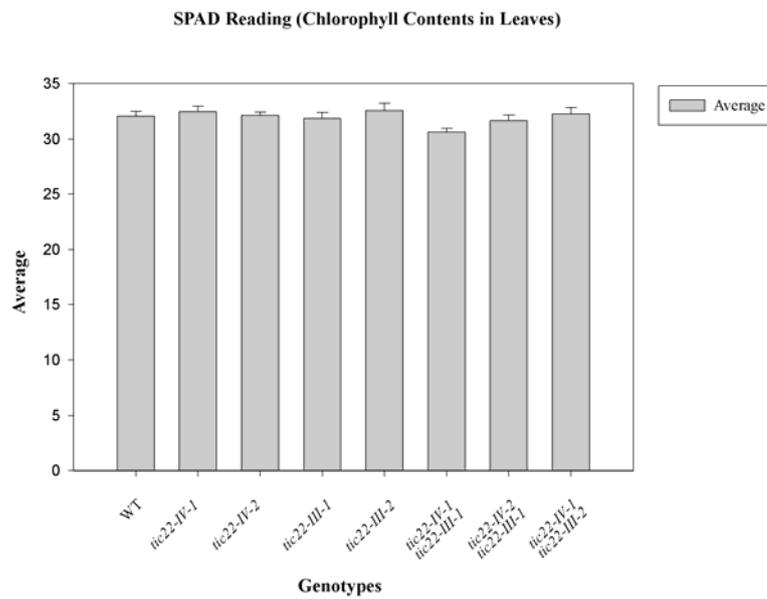


Figure 4.11. Chloroplast frequency size and distributions of 5-day-old plants.

(a) The cross sectional area (μm^2) measurement of chloroplasts from fully expanded leaves of wild type, single and double *tic22* (homologues) mutant. At least 50 chloroplasts were measured for each genotype. (b) The chloroplast shape (length/width) of the same plants. The values shown are means (\pm SE).

(a)



(b)

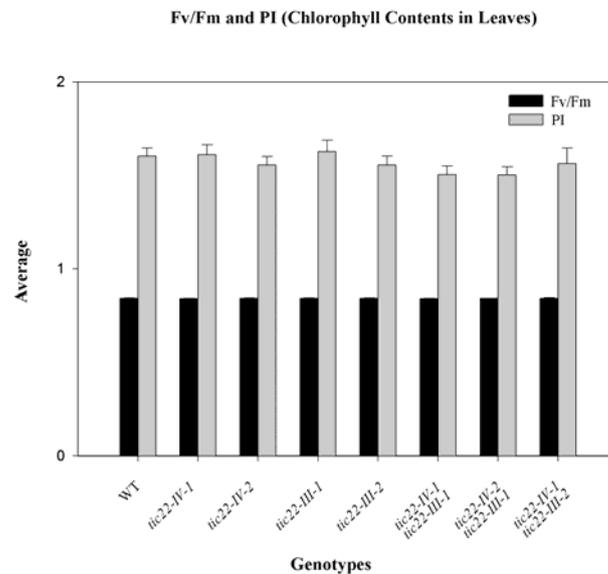


Figure 4.12. Phenotypic analysis of 30-days-old *tic22* single and double mutant plants.

(a) Chlorophyll concentrations in leaves of 30-day-old plants of the indicated genotypes were measured. The plants were grown *in vitro* for 14 days, and thereafter for an additional 16 days on soil. Values shown are means (\pm SE) derived from measurements of 10-15 different plants. Units are SPAD units.

(b) Analysis of photosynthesis in single- and double-mutant plants. Chlorophyll fluorescence measurements were used to estimate the photochemical efficiency of photosystem II (F_v/F_m) and PI the “Performance Index” (after 20 minutes dark adaptation). Measurements were done on fully-grown leaves from ten different 30-day-old plants per genotype; all plants were grown under identical conditions, as described in (a). Values shown are means (\pm SE) derived from analyses 10 individual plants.

4.3.7. Summary

This study showed *atTIC22-IV* and *atTIC22-III* knockout mutants individually do not have any effect on chloroplast development. However, when both genes are knocked out it seems that they have an important role at early stages of chloroplast development. Future study could experimentally prove in more detail that both *atTic22* proteins are membrane proteins localized to the periphery of the chloroplasts. Also the roles of both *atTic22-IV* and *atTic22-III* in protein translocation need to be tested directly.

Chapter 5

General Discussion

5.1. Preamble

Nuclear-encoded gene expression and chloroplast protein import are both believed to be integral parts of chloroplast development. In this work, I investigated the role and localization of both atTic20 (IEM) and atTic22 (IMS) in chloroplast biogenesis. Prior studies indicated that Tic20 and Tic22 are components of the protein import machinery, and are associated within TIC complexes (Kouranov *et al.*, 1998). However, their functions during protein import into the chloroplast need to be studied in more detail.

5.2. Tic20

Previously in *Arabidopsis* four *TIC20* genes were identified on the basis of similarity with the original isolate from pea (Bédard and Jarvis, 2005; Kalanon and McFadden, 2008). In this study, we have shown that these four homologues are expressed, that they are likely to be topologically similar to the pea protein, and that they are all targeted to the chloroplast envelope (Figure 3.4). The *Arabidopsis* and pea Tic20 proteins have been compared to those in other species to assess similarities and differences between the proteins. This phylogenetic analysis of a large number Tic20-related sequences revealed two evolutionarily conserved sub-classes of Tic20 proteins, termed Group 1 (characterized by atTic20-I and atTic20-IV) and Group 2 (characterized by atTic20-II and atTic20-V) (Figure 3.3). Analysis of the tree and of the corresponding sequence alignment revealed high sequence similarity of psTic20 to the Group 1 proteins (atTic20-I and atTic20-IV) and much less similarity to Group 2 proteins (atTic20-II and atTic20-V).

Genetic analysis, as shown in Figures 3.1a and 3.2, indicates that Group 1 proteins are more important in *Arabidopsis* development, as the loss of these proteins results in albinism or developmental arrest during gametophyte or embryo development. These defects are consistent with plastid dysfunction, and so may be caused mainly by defects in plastid protein import of nuclear-encoded proteins. Although the *tic20-I* mutants had severe albino phenotypes and are normally seedling lethal, these mutants are able to produce albino leaves on synthetic media supplemented with sucrose (5%).

The albino phenotype associated with the *tic20-I* T-DNA lines (Figure 3.7) is consistent with previous results (Teng *et al.*, 2006), and with the observation that

antisense-mediated down-regulation of *atTIC20-I* causes a greening defect (Chen *et al.*, 2002). Our genetic and complementation analyses also demonstrated conclusively that the two *Arabidopsis* Group 1 proteins share considerable functional redundancy (Figures 3.1a and 3.13). This redundancy is incomplete, however, as overexpression of *atTIC20-IV* could only mediate partial complementation of the *tic20-I-1* chloroplast defect. As it is shown in Chapter 3, only *atTIC20-I* knockout mutants have a severe phenotype (albinism) in isolation; this shows that that protein has an important role in chloroplast and photosynthetic development, presumably as a channel protein. It is interesting that the *tic20-I* mutant phenotype is not even more severe (e.g., like the mutant for the outer membrane channel protein, atToc75-III, which is embryo lethal (Baldwin *et al.*, 2005). This may be due to the presence of other Tic20 isoforms (for example *atTIC20-IV*; indeed, our double mutants confirm such a partial redundancy relationship). It is possible that the different Tic20 isoforms have somewhat different substrate specificities, perhaps like the TOC receptor proteins (Jarvis, 2008a). The atTic20-IV and/or the Group 2 proteins (atTic20-II and atTic20-V) in *Arabidopsis* may be responsible for the import of non-photosynthetic or housekeeping proteins.

As indicated in Chapter 3, it appears that all four *Arabidopsis* *TIC20* genes are expressed throughout development (Figure 3.5). Nonetheless, the *atTIC20-I* and *atTIC20-IV* genes do exhibit quite different patterns of expression, suggesting that the former is relatively more important for photosynthetic development, and that the latter acts primarily during non-photosynthetic growth and seed development. The high expression of *atTIC20-IV* in seeds is consistent with the observation that its inactivation, in conjunction with *tic20-I* mutations, causes developmental arrest during either female gametogenesis or embryogenesis. The fact that these two genes exhibit distinct expression patterns, biased in favour of either photosynthetic or non-photosynthetic growth, parallels observations made in relation to the TOC receptors, Toc159 and Toc33 (Jarvis *et al.*, 1998; Bauer *et al.*, 2000; Gutensohn *et al.*, 2000; Kubis *et al.*, 2003; Kubis *et al.*, 2004). Detailed analyses of different isoforms of these receptors has indicated that the dominant one in each case (atToc159 and atToc33, respectively) is specialized for the import of highly-abundant, photosynthesis-related preproteins, and that the lesser isoforms (atToc132, atToc120 and atToc34) preferentially mediate import of low-abundance, housekeeping preproteins. In this regard, it is noteworthy that the analysis of *tic20-I* mutant plants using an *in vivo* import assay revealed a strong effect

on the import of a photosynthetic precursor, and little effect on that of a housekeeping protein (Kikuchi *et al.*, 2009).

One hypothesis suggested to account for the sub-functionalization of TOC receptor isoforms is that it prevents potentially destructive competition effects between preproteins; in the absence of such client-specific receptor complexes, the bulk-flow of photosynthetic precursors might interfere with the import of other preproteins (Kessler and Schnell, 2006a; Jarvis, 2008a). It has been suggested that the different import pathways converge at the TIC machinery, based on observations that some TIC components (e.g., Tic110) have not undergone similar sub-functionalization in *Arabidopsis* (Kovacheva *et al.*, 2005). However, recent evidence suggests that there may be at least two different TIC complexes, and that the one containing Tic110 acts downstream of that containing Tic20 (Kikuchi *et al.*, 2009). If this is in fact the case, it is possible that as mentioned previously client-specific import pathways do extend to the level of the inner membrane (at Tic20), and that convergence only happens later at Tic110. In other words, there may be distinct Tic20-containing channel complexes (e.g., containing either atTic20-I or atTic20-IV), but just a single Tic110-containing chaperone complex.

Several components of the chloroplast protein transport machinery are known to be essential for embryo viability (Baldwin *et al.*, 2005; Inaba *et al.*, 2005; Kovacheva *et al.*, 2005; Hust and Gutensohn, 2006; Patel *et al.*, 2008), but there have been relatively few reports of lesions in chloroplast proteins leading to gametophyte arrest (which is what we observed in the case of the *tic20-I tic20-IV* double mutants). Examples are the *gpt1* mutations that affect the glucose 6-phosphate/phosphate translocator (GPT) of the chloroplast inner envelope membrane (Niewiadomski *et al.*, 2005). However, in contrast with the Group 1 Tic20 double mutants described here, the *gpt1* mutants exhibited defective transmission through both male and female gametes; these effects are attributed to a reduced supply of reducing equivalents via the oxidative pentose phosphate pathway, which in turn affects fatty acid synthesis leading to defective membrane biogenesis. That the gametophytic defects seen in *tic20-I tic20-IV* are somewhat less severe may indicate that the Tic20 complex is not essential for the biogenesis of inner membrane carrier proteins such as GPT, or that pre-meiotically synthesized Tic20 persists long enough to mediate sufficient GPT biogenesis.

Another example of plastid-linked defective gametogenesis occurs in the *Arabidopsis* Hsp93 double mutant (Kovacheva *et al.*, 2007). Transmission of the *hsp93-*

V-1 hsp93-III-2 double mutation through female gametes was reduced to 46.8%, and as a consequence a significant number of failed ovules were observed in the siliques of heterozygous double-mutant plants. These results are similar to those reported here, although the female transmission efficiency defect was much stronger in the case of the *tic20-I tic20-IV* double mutation (at 34.4%), and as a consequence a much greater proportion of defective reproductive structures (mostly failed ovules) was observed in the mutant siliques. This severity of the Group 1 Tic20 double-mutant phenotype during reproductive growth (Figure 3.9b) parallels that seen in single-mutant *tic20-I* plants (which are even more sick than the *ppi2* albino; Figure 3.9a), and suggests that the Tic20 proteins play a central and crucial role in the import mechanism (e.g., translocation channel formation).

The roles of the Group 2 Tic20 proteins remain uncertain, as no mutant phenotypes could be detected in any of the mutant genotypes analysed (Figure 3.12a). Nonetheless, the observations that these genes are expressed at relatively high levels (Figure 3.6) and have been conserved over millions of years of evolution in many diverse species (Figure 3.2) together suggest that they do perform some important role. One possibility is that these proteins are somehow analogous to Tim17 in mitochondria; the role of this protein is uncertain, although it has been proposed to regulate the channel formed by the related protein, Tim23 (Neupert and Herrmann, 2007). In contrast with the Group 2 Tic20 proteins, however, Tim17 is essential for viability in yeast. It is likely that the roles of atTic20-II and atTic20-V become critical under non-standard growth conditions (e.g., various stresses) that were not tested during the course of this study.

5.3. Tic22

In *Arabidopsis* two *TIC22* genes were identified on the basis of similarity with the original isolate from pea (Bédard and Jarvis, 2005). In this study, we have showed that these two homologues are expressed, that they are likely to be topologically similar to the pea protein, and that they are both likely targeted to the chloroplast envelope (Figure 4.1). Prior study indicated that Tic22 is located in the chloroplast intermembrane space (Kouranov *et al.*, 1998). Recognition of the precursor proteins believed to be mediated by receptor proteins at the outer membrane proteins as stated previously (Toc34 and

Toc159). The preproteins then pass through the TOC channel protein, atToc75-III, and then on emerging from that channel it is suggested that they are recognized by Tic22 en route to the TIC machinery (e.g., Tic20, Tic110, etc.).

The *Arabidopsis* and pea Tic22 proteins have been compared with those in other species to assess similarities and differences between these proteins. My phylogenetic analysis of *Arabidopsis* Tic22-related sequences revealed two evolutionarily conserved sub-classes of Tic22 proteins, termed Group 1 and Group 2 (Figure 4.3). Prior study revealed that the *Arabidopsis* genome has undergone duplication and reshuffling (Blanc *et al.*, 2000); these modifications may have led to the divergence of the atTic22 family, through duplication and mutations. However, my analysis revealed that the two groups of Tic22-related proteins actually diverged early in land plant evolution. As the pea genome sequence is not completed, consequently it is not possible to predict how many pea Tic22 homologues exist. Comparison study of *Arabidopsis* Tic22 genes also showed that *atTIC22-III* has lost a single intron (Figure 4.9a), by comparison with the *atTIC22-IV* gene. Previously, the loss of a single intron in plants has been observed on a number of occasions (Huang *et al.*, 1990; Kumar and Trick, 1993; Häger *et al.*, 1996; Drouin and Moniz de Sa, 1997). One possible mechanism by which *atTIC22-III* might have lost an intron would be through the reverse transcription of mRNA to produce a cDNA copy of gene, which could then partially replace the endogenous copy via homologous recombination (Baltimore, 1985; Derr *et al.*, 1991).

My genetic analyses of *atTIC22-IV* and *atTIC22-III* (Figure 4.10a and b) showed that the single knockout mutants individually do not have any effect on chloroplast development. However, when both of these genes are knocked out it seems that they have an important role at early stages of chloroplast development. As plants grow older the chlorophyll content of these double mutants increases, and by day 20 there is no reduction in chlorophyll concentration observed, relative to wild type. This observation may indicate that *atTIC22* homologues are expressed at particularly high levels during later development (Figures 4.10a and 4.12). Although the *atTIC22-IV* gene expresses at relatively higher levels than *atTIC22-III*, it seems that the expression of *atTIC22-IV* is not particularly high (Figure 4.6). Moreover, neither of the genes is expressed at much higher level during an early development compared with later stages. Another potential explanation for double-mutant phenotype is that the role of Tic22 (-IV and -III) is only critical during early stages of chloroplast biogenesis. In

summary, my data indicated that both these genes are not essential for chloroplast biogenesis throughout development.

Prior studies have shown that the protein import of Tic22 depends on its N-terminal presequence (Kouranov *et al.*, 1999). The Tic22 protein import study revealed that the imported of preTic22 appears to be a slow event in the import process (Kouranov *et al.*, 1999). They have shown that the huge amount of preTic22-imported remains unprocessed and the strength of preTic22 associate with the inner membrane suggest that cleavage of the precursor was not strongly joined to translocation across the outer membrane or binding to the inner membrane (Kouranov *et al.*, 1999). It was supposed that Tic22 is processed in the intermembrane space by unknown peptidase. It was also proposed that the targeting pathway to the intermembrane space is unrelated to other known chloroplast targeting pathways which proceed at the envelope membranes (Kouranov *et al.*, 1999). This conclusion was partly based on the fact that Tic22 requires only low levels of ATP in order to reach its final destination (Kouranov *et al.*, 1999).

To investigate whether the *Arabidopsis* Tic22 homologues are chloroplast envelope proteins, the YFP and chloroplast import studies were conducted. In general, these studies supported the conclusion that the two components are chloroplast membrane proteins. Having said that, the Tic22-YFP fluorescence data were not as well defined as those seen for the Tic20-YFP fusions (Chapter 3), and this may be due in part to the unique targeting properties of IMS proteins like Tic22 mentioned above. It was shown in Chapter 4 that atTic22-IV is the main homologue of psTic22; therefore, this protein should behave the same as psTic22 in terms of its import into chloroplasts. Because atTic22-III is less similar to psTic22, import of atTic22-III was excluded from this study. The chloroplast protein import of atTic22-IV suggested that this protein is a chloroplast membrane protein, although the results were not conclusive and need to be repeated. In future, the chloroplast import into the intermembrane space of atTic22 homologues needs to be investigating in more detail. Future characterization of these proteins in terms of their presequence processing sites, as well as their interaction with TOC complexes, should enable us to obtain a clearer picture of how they are targeted.

As there is lack of data related to chloroplast import into the intermembrane space, other organelles like mitochondria could possibly assist us to improve our understanding about the mechanism of protein import into chloroplasts. Like chloroplasts, mitochondria possess two “envelope” membranes. Interestingly,

techniques by which to remove the mitochondrial outer envelope membrane have been developed and successfully used. In mitochondria, the outer membrane component could be burst or dissolved, by osmotic shock treatment (Daum *et al.*, 1982) or by digitonin treatment (Hartl *et al.*, 1986), respectively, in order to produce so-called mitoplasts. These two methods have effectively been used to study the localization and topology of mitochondrial envelope membrane proteins. Unlike mitochondria, there are not yet any similar techniques by which to selectively remove the outer membrane of the envelope of chloroplasts. But such methods, if developed, may aid us in the future.

5.4. Conclusion

In conclusion, chloroplasts continue to be a mysterious feature of the plant cell. The aim of this research was to provide a better understanding about the role in development and localization of atTic20 and atTic22 homologues in chloroplasts. The next priority is to study their interactions with other TOC or TIC proteins in chloroplast supercomplexes. As it was mentioned before, (in Chapter 1), since our knowledge of the TIC complex is very limited, further investigation is required to determine if there is any interaction between different TIC components, for example between Tic20 and Tic22. This possibly might point towards the functional cooperation between these two proteins during protein import. Genetic mutation analyses can have a major impact on protein import studies in chloroplasts; this approach has many great advantages (foremost of which is physiological relevance and the capacity to study processes *in vivo*), and in the future it should be combined more with complementary approaches such as physiological and biochemical studies to further investigate plant growth and chloroplast biogenesis.

Appendix 3.1.

Target P v1.1 prediction results

Sequence atTic20-I cDNA, AK117165-seq

Number of query sequences: 1
Cleavage site predictions not included.
Using PLANT networks.

Name	Len	cTP	mTP	SP	other	Loc	RC	TPlen
Sequence	274	0.519	0.259	0.001	0.065	C	4	65
cutoff		0.000	0.000	0.000	0.000			

Target P v1.1 prediction results

Sequence atTic20-IV cDNA, BX828045-seq

Number of query sequences: 1
Cleavage site predictions included.
Using PLANT networks.

Name	Len	cTP	mTP	SP	other	Loc	RC	TPlen
Sequence	284	0.611	0.324	0.000	0.112	C	4	48
cutoff		0.000	0.000	0.000	0.000			

Target P v1.1 prediction results

Sequence atTic20-II cDNA, BX821678-seq

Number of query sequences: 1
Cleavage site predictions not included.
Using PLANT networks.

Name	Len	cTP	mTP	SP	other	Loc	RC	TPlen
Sequence	208	0.905	0.324	0.004	0.025	C	3	49
cutoff		0.000	0.000	0.000	0.000			

Target P v1.1 prediction results

Sequence atTic20-V cDNA - AY087311-seq

Number of query sequences: 1
Cleavage site predictions not included.
Using PLANT networks.

Name	Len	cTP	mTP	SP	other	Loc	RC	TPlen
Sequence	209	0.456	0.103	0.193	0.102	C	4	49
cutoff		0.000	0.000	0.000	0.000			

TargetP 1.1 Server - prediction results

Sequence psTic20

Number of query sequences: 1
Cleavage site predictions not included.
Using PLANT networks.

Name	Len	cTP	mTP	SP	other	Loc	RC	TPlen
Sequence	253	0.525	0.052	0.024	0.127	C	4	42
cutoff		0.000	0.000	0.000	0.000			

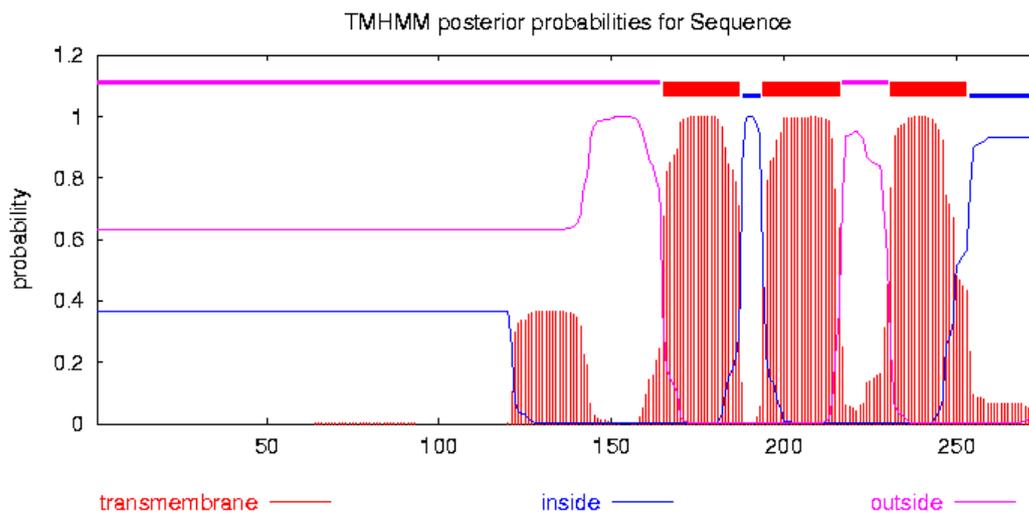
Appendix 3.2.

atTic20-I

TMHMM result

[HELP](#) with output formats

```
# Sequence Length: 274
# Sequence Number of predicted TMHs: 3
# Sequence Exp number of AAs in TMHs: 74.09372
# Sequence Exp number, first 60 AAs: 0.00137
# Sequence Total prob of N-in: 0.36796
Sequence      TMHMM2.0      outside      1      164
Sequence      TMHMM2.0      TMhelix     165     187
Sequence      TMHMM2.0      inside      188     193
Sequence      TMHMM2.0      TMhelix     194     216
Sequence      TMHMM2.0      outside     217     230
Sequence      TMHMM2.0      TMhelix     231     253
Sequence      TMHMM2.0      inside      254     274
```



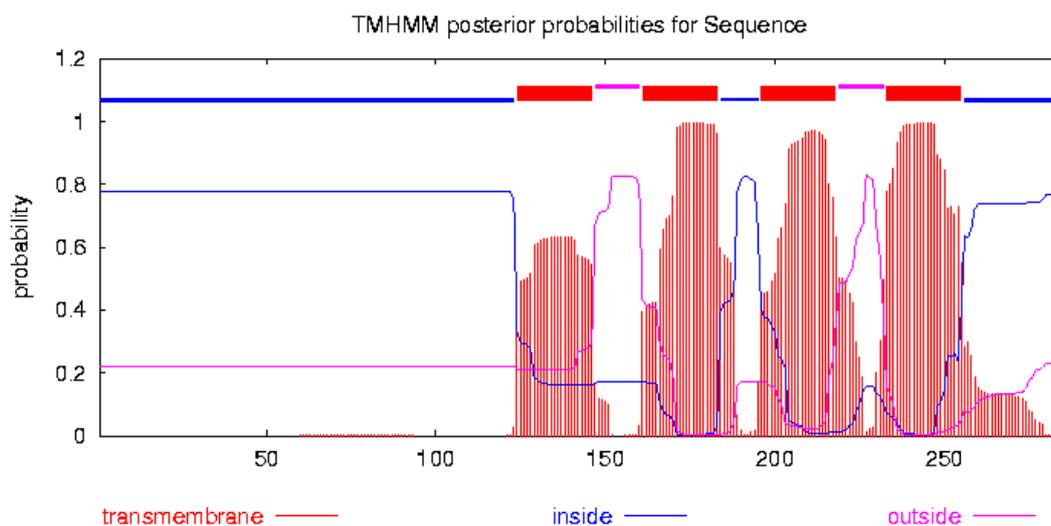
[plot](#) in postscript, [script](#) for making the plot in gnuplot, [data](#) for plot

atTic20-IV

TMHMM result

[HELP](#) with output formats

```
# Sequence Length: 284
# Sequence Number of predicted TMHs: 4
# Sequence Exp number of AAs in TMHs: 81.23357
# Sequence Exp number, first 60 AAs: 0.00119
# Sequence Total prob of N-in: 0.77822
Sequence      TMHMM2.0      inside      1      123
Sequence      TMHMM2.0      TMhelix    124    146
Sequence      TMHMM2.0      outside    147    160
Sequence      TMHMM2.0      TMhelix    161    183
Sequence      TMHMM2.0      inside    184    195
Sequence      TMHMM2.0      TMhelix    196    218
Sequence      TMHMM2.0      outside    219    232
Sequence      TMHMM2.0      TMhelix    233    255
Sequence      TMHMM2.0      inside    256    284
```



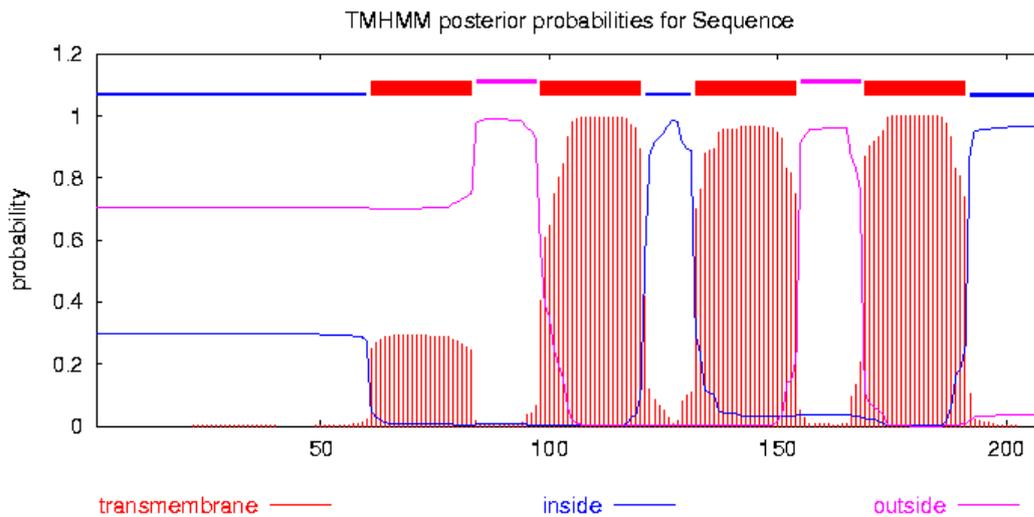
[plot](#) in postscript, [script](#) for making the plot in gnuplot, [data](#) for plot

atTic20-II

TMHMM result

[HELP](#) with output formats

```
# Sequence Length: 208
# Sequence Number of predicted TMHs: 4
# Sequence Exp number of AAs in TMHs: 72.06447
# Sequence Exp number, first 60 AAs: 0.07078
# Sequence Total prob of N-in: 0.29628
Sequence      TMHMM2.0      inside      1      60
Sequence      TMHMM2.0      TMhelix     61     83
Sequence      TMHMM2.0      outside     84     97
Sequence      TMHMM2.0      TMhelix     98    120
Sequence      TMHMM2.0      inside    121    131
Sequence      TMHMM2.0      TMhelix    132    154
Sequence      TMHMM2.0      outside    155    168
Sequence      TMHMM2.0      TMhelix    169    191
Sequence      TMHMM2.0      inside    192    208
```



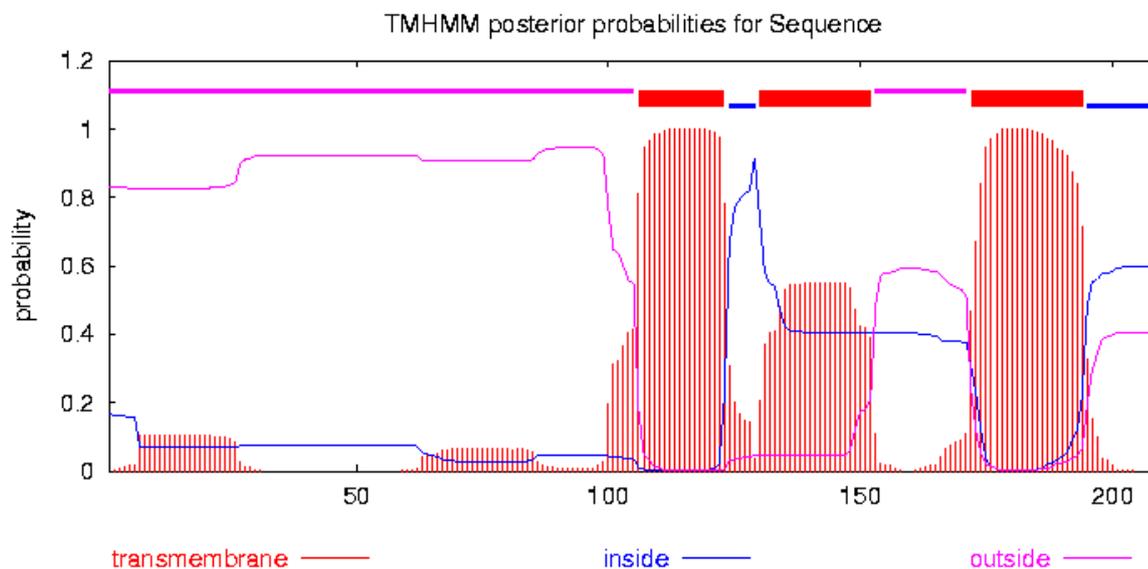
[plot](#) in postscript, [script](#) for making the plot in gnuplot, [data](#) for plot

atTic20-V

TMHMM result

[HELP](#) with output formats

```
# Sequence Length: 209
# Sequence Number of predicted TMHs: 3
# Sequence Exp number of AAs in TMHs: 57.77705
# Sequence Exp number, first 60 AAs: 2.13687
# Sequence Total prob of N-in: 0.16709
Sequence      TMHMM2.0      outside      1      105
Sequence      TMHMM2.0      TMhelix     106     123
Sequence      TMHMM2.0      inside      124     129
Sequence      TMHMM2.0      TMhelix     130     152
Sequence      TMHMM2.0      outside     153     171
Sequence      TMHMM2.0      TMhelix     172     194
Sequence      TMHMM2.0      inside      195     209
```



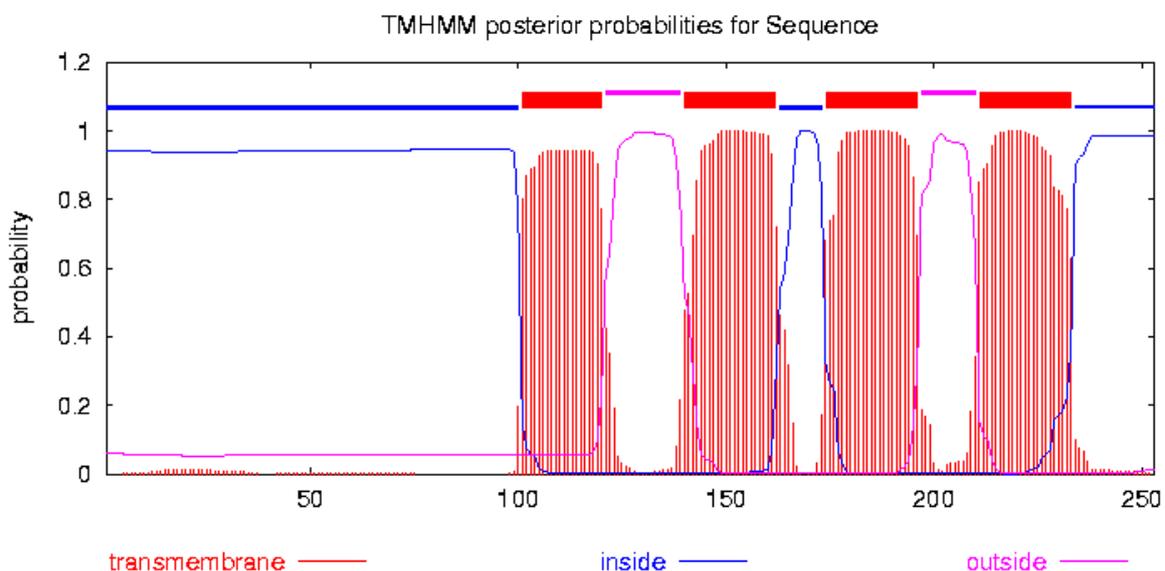
psTic20

TMHMM result

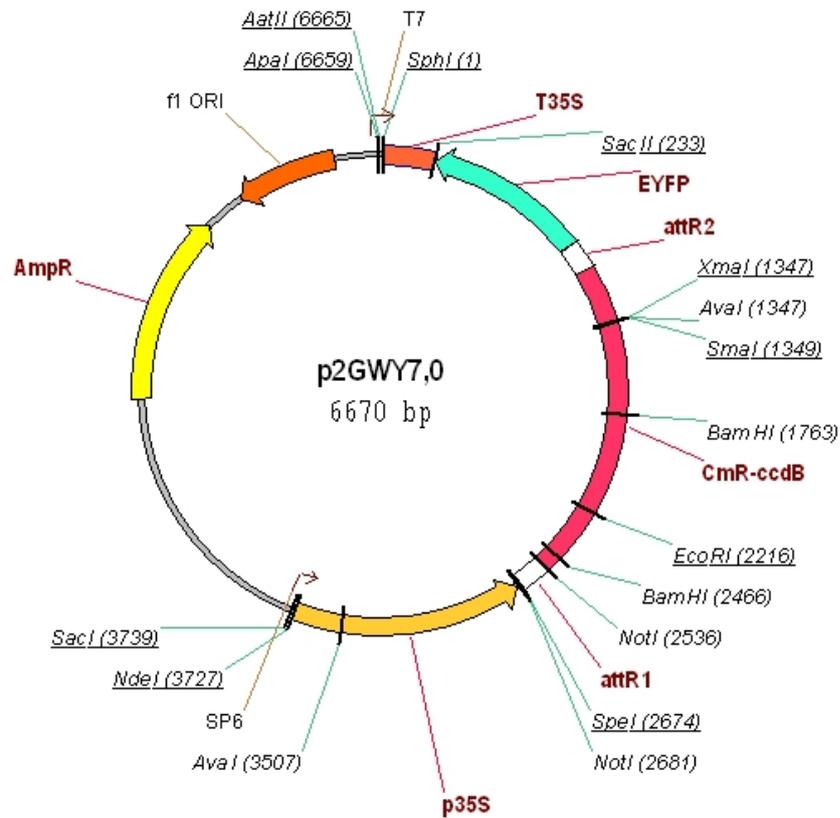
[HELP](#) with output formats

```
# Sequence Length: 253
# Sequence Number of predicted TMHs: 4
# Sequence Exp number of AAs in TMHs: 87.18858
# Sequence Exp number, first 60 AAs: 0.25932
# Sequence Total prob of N-in: 0.94147
Sequence      TMHMM2.0      inside      1    100
Sequence      TMHMM2.0      TMhelix    101  120
Sequence      TMHMM2.0      outside    121  139
Sequence      TMHMM2.0      TMhelix    140  162
Sequence      TMHMM2.0      inside    163  173
Sequence      TMHMM2.0      TMhelix    174  196
Sequence      TMHMM2.0      outside    197  210
Sequence      TMHMM2.0      TMhelix    211  233
Sequence      TMHMM2.0      inside    234  253
```

[plot](#) in postscript, [script](#) for making the plot in gnuplot, [data](#) for plot

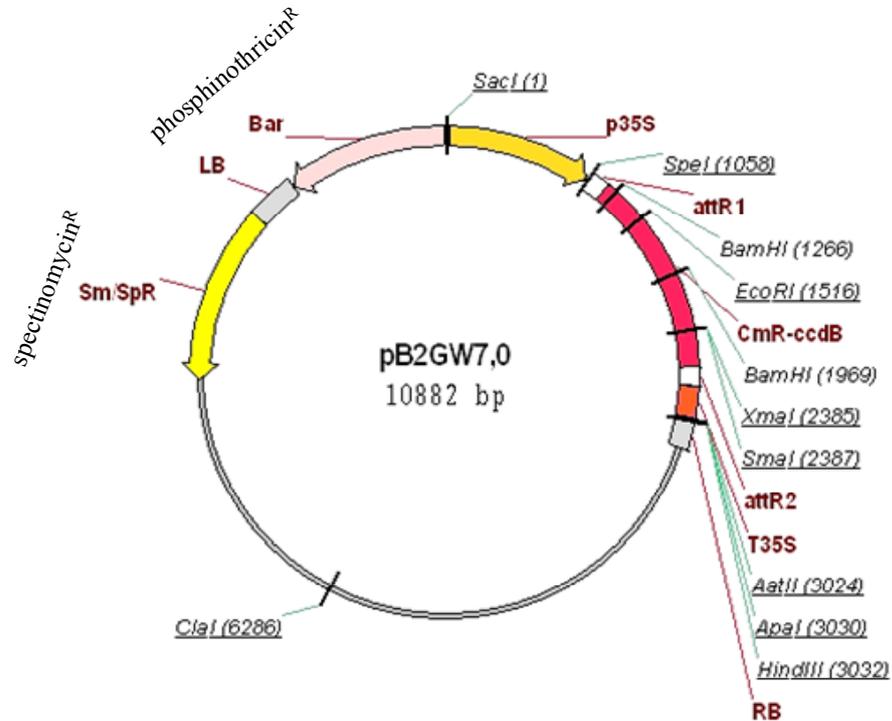


Appendix 3.3.



Supplementary Figure 3.1. Gateway destination vector used for YFP mediated plant transformation (Karimi *et al.*, 2005), with the Cm^R (chloramphenicol) antibiotic to select gene in the GATEWAY cassette and Amp^R (ampicillin) selective marker for bacterial. EYFP is an enhanced YFP, attR1 and attR2 are the recombination gateway cassette. 35S is the cauliflower mosaic virus 35S promoter and terminator.

Appendix 3.4.



Supplementary Figure 3.2. Gateway destination vectors used for overexpression Agrobacterium-mediated plant transformation (Karimi *et al.*, 2005). The Cm^R (chloramphenicol) antibiotic to select gene in the GATEWAY cassette and sm/Sp^R (spectinomycin) selective marker for bacterial and Bar is phosphinothricine. The attR1 and attR2 are the recombination gateway cassette. 35S is the cauliflower mosaic virus 35S promoter and terminator.

Appendix 4.1.

Target P v1.1 prediction results

Sequence atTic22-IV cDNA

Number of query sequences: 1
Cleavage site predictions not included.
Using PLANT networks.

Name	Len	cTP	mTP	SP	other	Loc	RC	TPlen
Sequence	268	0.588	0.078	0.004	0.106	C	3	59
cutoff		0.000	0.000	0.000	0.000			

Target P v1.1 prediction results

Sequence atTic22-III cDNA NM_11275.2-seq

Number of query sequences: 1
Cleavage site predictions not included.
Using PLANT networks.

Name	Len	cTP	mTP	SP	other	Loc	RC	TPlen
Sequence	313	0.927	0.046	0.007	0.129	C	2	96
cutoff		0.000	0.000	0.000	0.000			

Target P v1.1 prediction results

Sequence psTic22

targetp v1.1 prediction results #####
Number of query sequences: 1
Cleavage site predictions included.
Using PLANT networks.

Name	Len	cTP	mTP	SP	other	Loc	RC	TPlen
Sequence	252	0.421	0.511	0.005	0.129	M	5	28
cutoff		0.000	0.000	0.000	0.000			

References

- Abe, Y., Shodai, T., Muto, T., Mihara, K., Torii, H., Nishikawa, S., Endo, T., and Kohda, D.** (2000). Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. *Cell* **100**, 551-560.
- Akita, M., Nielsen, E., and Keegstra, K.** (1997). Identification of protein transport complexes in the chloroplastic envelope membranes via chemical cross-linking. *J. Cell Biol.* **136**, 983-994.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J.** (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J.** (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402.
- America, T., Hageman, J., Guera, A., Rook, F., Archer, K., Keegstra, K., and Weisbeek, P.** (1994). Methotrexate does not block import of a DHFR fusion protein into chloroplasts. *Plant Mol. Biol.* **24**, 283-294.
- Archibald, J., and Keeling, P.** (2004). Progress and Problems in the Origin and Evolution of Plastids. In Sapp J (ed) *Microbial Evolution. Concepts and Controversies*. Oxford University Press, Oxford, in press.
- Aronsson, H., and Jarvis, P.** (2002). A simple method for isolating import-competent *Arabidopsis* chloroplasts. *FEBS Lett.* **529**, 215-220.
- Aronsson, H., Combe, J., and Jarvis, P.** (2003). Unusual nucleotide-binding properties of the chloroplast protein import receptor, atToc33. *FEBS Lett.* **544**, 79-85.
- Aronsson, H., Boij, P., Patel, R., Wardle, A., Töpel, M., and Jarvis, P.** (2007). Toc64/OEP64 is not essential for the efficient import of proteins into chloroplasts in *Arabidopsis thaliana*. *Plant J.* **52**, 53-68.
- Aronsson, H., Schottler, M.A., Kelly, A.A., Sundqvist, C., Dormann, P., Karim, S., and Jarvis, P.** (2008). Monogalactosyldiacylglycerol deficiency in *Arabidopsis thaliana* affects pigment composition in the prolamellar body and impairs thylakoid membrane energization and photoprotection in leaves. *Plant Physiol.*
- Baginsky, S., and Gruissem, W.** (2004). Chloroplast proteomics: potentials and challenges. *J. Exp. Bot.* **55**, 1213-1220.
- Bähring, R., Milligan, C.J., Vardanyan, V., Engeland, B., Young, B.A., Dannenberg, J., Waldschutz, R., Edwards, J.P., Wray, D., and Pongs, O.** (2001). Coupling of voltage-dependent potassium channel inactivation and oxidoreductase active site of K ν beta subunits. *J. Biol. Chem.* **276**, 22923-22929.
- Baldwin, A., Wardle, A., Patel, R., Dudley, P., Park, S.K., Twell, D., Inoue, K., and Jarvis, P.** (2005). A molecular-genetic study of the *Arabidopsis* Toc75 gene family. *Plant Physiol.* **138**, 1-19.
- Balsera, M., Stengel, A., Soll, J., and Bölder, B.** (2007). Tic62: a protein family from metabolism to protein translocation. *BMC Evol. Biol.* **7**, 43.
- Balsera, M., Goetze, T.A., Kovács-Bogdán, E., Schürmann, P., Wagner, R., Buchanan, B.B., Soll, J., and Bölder, B.** (2009). Characterization of Tic110, a channel-forming protein at the inner envelope membrane of chloroplasts, unveils a response to Ca²⁺ and a stromal regulatory disulfide bridge. *J. Biol. Chem.* **284**, 2603-2616.
- Baltimore, D.** (1985). Retroviruses and retrotransposons: the role of reverse transcription in shaping the eukaryotic genome. *Cell* **40**, 481-482.

- Bauer, J., Chen, K., Hiltbunner, A., Wehrli, E., Eugster, M., Schnell, D., and Kessler, F.** (2000). The major protein import receptor of plastids is essential for chloroplast biogenesis. *Nature* **403**, 203-207.
- Becker, T., Jelic, M., Vojta, A., Radunz, A., Soll, J., and Schleiff, E.** (2004a). Preprotein recognition by the Toc complex. *EMBO J.* **23**, 520-530.
- Becker, T., Hritz, J., Vogel, M., Caliebe, A., Bukau, B., Soll, J., and Schleiff, E.** (2004b). Toc12, a novel subunit of the intermembrane space preprotein translocon of chloroplasts. *Mol. Biol. Cell* **15**, 5130-5144.
- Bédard, J., and Jarvis, P.** (2005). Recognition and envelope translocation of chloroplast preproteins. *J. Exp. Bot.* **56**, 2287-2320.
- Bédard, J., Kubis, S., Bimanadham, S., and Jarvis, P.** (2007). Functional similarity between the chloroplast translocon component, Tic40, and the human co-chaperone, Hsp70-interacting protein (Hip). *J. Biol. Chem.* **282**, 21404-21414.
- Bendich, A.J.** (2004). Circular chloroplast chromosomes: the grand illusion. *Plant Cell* **16**, 1661-1666.
- Besendahl, A., Qiu, Y.L., Lee, J., Palmer, J.D., and Bhattacharya, D.** (2000). The cyanobacterial origin and vertical transmission of the plastid tRNA(Leu) group-I intron. *Curr Genet* **37**, 12-23.
- Bhushan, S., Kuhn, C., Berglund, A.K., Roth, C., and Glaser, E.** (2006). The role of the N-terminal domain of chloroplast targeting peptides in organellar protein import and miss-sorting. *FEBS Lett.* **580**, 3966-3972.
- Birnboim, H.C., and Doly, J.** (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* **7**, 1513-1523.
- Blanc, G., Barakat, A., Guyot, R., Cooke, R., and Delseny, M.** (2000). Extensive duplication and reshuffling in the *Arabidopsis* genome. *Plant Cell* **12**, 1093-1101.
- Blobel, G., and Sabatini, D.D.** (1971). Ribosome-membrane interactions in eukaryotic cells. In *Biomembranes*, L.A. Manson, ed (New York: Plenum), pp. 193-195.
- Bogsch, E.G., Sargent, F., Stanley, N.R., Berks, B.C., Robinson, C., and Palmer, T.** (1998). An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. *J. Biol. Chem.* **273**, 18003-18006.
- Bohnert, M., Pfanner, N., and van der Laan, M.** (2007). A dynamic machinery for import of mitochondrial precursor proteins. *FEBS Lett.* **581**, 2802-2810.
- Bourne, H.R., Sanders, D.A., and McCormick, F.** (1991). The GTPase superfamily: conserved structure and molecular mechanism. *Nature* **349**, 117-127.
- Bouvier, F., Backhaus, R., and Camara, B.** (1998). Induction and Control of Chromoplast-specific Carotenoid Genes by Oxidative Stress. *Nature* **273**, 30651-30659.
- Bramhill, D.** (1997). Bacterial cell division. *Annu Rev Cell Dev Biol* **13**, 395-424.
- Brix, J., Dietmeier, K., and Pfanner, N.** (1997). Differential recognition of preproteins by the purified cytosolic domains of the mitochondrial import receptors Tom20, Tom22, and Tom70. *J. Biol. Chem.* **272**, 20730-20735.
- Bruce, B.D.** (1998). The role of lipids in plastid protein transport. *Plant Mol Biol* **38**, 223-246.
- Bruce, B.D.** (2000). Chloroplast transit peptides: structure, function and evolution. *Trends Cell Biol.* **10**, 440-447.
- Bruce, B.D.** (2001). The paradox of plastid transit peptides: conservation of function despite divergence in primary structure. *Biochim. Biophys. Acta* **1541**, 2-21.

- Buchanan, B., Gruissem, W., and Jones, R.** (2000). *Biochemistry and Molecular Biology of Plants*. Rockville, Maryland: American Society of Plant Physiologists.
- Caliebe, A., Grimm, R., Kaiser, G., Lubeck, J., Soll, J., and Heins, L.** (1997). The chloroplastic protein import machinery contains a Rieske-type iron-sulfur cluster and a mononuclear iron-binding protein. *Embo J* **16**, 7342-7350.
- Chen, K., Chen, X., and Schnell, D.J.** (2000). Initial binding of preproteins involving the Toc159 receptor can be bypassed during protein import into chloroplasts. *Plant Physiol.* **122**, 813-822.
- Chen, K.Y., and Li, H.M.** (2007). Precursor binding to an 880-kDa Toc complex as an early step during active import of protein into chloroplasts. *Plant J.* **49**, 149-158.
- Chen, L.J., and Li, H.M.** (1998). A mutant deficient in the plastid lipid DGD is defective in protein import into chloroplasts. *Plant J* **16**, 33-39.
- Chen, X., Smith, M.D., Fitzpatrick, L., and Schnell, D.J.** (2002). In vivo analysis of the role of atTic20 in protein import into chloroplasts. *Plant Cell* **14**, 641-654.
- Chew, O., and Whelan, J.** (2004). Just read the message: a model for sorting of proteins between mitochondria and chloroplasts. *Trends Plant Sci.* **9**, 318-319.
- Chew, O., Lister, R., Qbadou, S., Heazlewood, J.L., Soll, J., Schleiff, E., Millar, A.H., and Whelan, J.** (2004). A plant outer mitochondrial membrane protein with high amino acid sequence identity to a chloroplast protein import receptor. *FEBS Lett.* **557**, 109-114.
- Chigri, F., Hörmann, F., Stamp, A., Stammers, D.K., Bölder, B., Soll, J., and Vothknecht, U.C.** (2006). Calcium regulation of chloroplast protein translocation is mediated by calmodulin binding to Tic32. *Proc. Natl. Acad. Sci. USA* **103**, 16051-16056.
- Chou, M.L., Chu, C.C., Chen, L.J., Akita, M., and Li, H.M.** (2006). Stimulation of transit-peptide release and ATP hydrolysis by a cochaperone during protein import into chloroplasts. *J. Cell Biol.* **175**, 893-900.
- Chou, M.L., Fitzpatrick, L.M., Tu, S.L., Budziszewski, G., Potter-Lewis, S., Akita, M., Levin, J.Z., Keegstra, K., and Li, H.M.** (2003). Tic40, a membrane-anchored co-chaperone homolog in the chloroplast protein translocon. *EMBO J.* **22**, 2970-2980.
- Clark, S.A., and Theg, S.M.** (1997). A folded protein can be transported across the chloroplast envelope and thylakoid membranes. *Mol. Biol. Cell* **8**, 923-934.
- Clegg, M.T., Gaut, B.S., Learn, G.H., Jr., and Morton, B.R.** (1994). Rates and patterns of chloroplast DNA evolution. *Proc Natl Acad Sci U S A* **91**, 6795-6801.
- Cline, K., and Henry, R.** (1996). Import and routing of nucleus-encoded chloroplast proteins. *Annu. Rev. Cell. Dev. Biol.* **12**, 1-26.
- Cline, K., Werner-Washburne, M., Andrews, J., and Keegstra, K.** (1984). Thermolysin is a suitable protease for probing the surface of intact pea chloroplasts. *Plant Physiol* **75**, 675-678.
- Clough, S.J., and Bent, A.F.** (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735-743.
- Constan, D., Patel, R., Keegstra, K., and Jarvis, P.** (2004). An outer envelope membrane component of the plastid protein import apparatus plays an essential role in *Arabidopsis*. *Plant J.* **38**, 93-106.
- Dabney-Smith, C., van Den Wijngaard, P.W., Treece, Y., Vredenberg, W.J., and Bruce, B.D.** (1999). The C terminus of a chloroplast precursor modulates its

- interaction with the translocation apparatus and PIRAC. *J Biol Chem* **274**, 32351-32359.
- Daum, G., Gasser, S.M., and Schatz, G.** (1982). Import of proteins into mitochondria. Energy-dependent, two-step processing of the intermembrane space enzyme cytochrome b2 by isolated yeast mitochondria. *J Biol Chem* **257**, 13075-13080.
- Davila-Aponte, J.A., Inoue, K., and Keegstra, K.** (2003). Two chloroplastic protein translocation components, Tic110 and Toc75, are conserved in different plastid types from multiple plant species. *Plant Mol. Biol.* **51**, 175-181.
- De Boer, A.D., and Weisbeek, P.J.** (1991). Chloroplast protein topogenesis: import, sorting and assembly. *Biochim Biophys Acta* **1071**, 221-253.
- Della-Cioppa, G., Bauer, S.C., Klein, B.K., Shah, D.M., Fraley, R.T., and Kishore, G.M.** (1986). Translocation of the precursor of 5-enolpyruvylshikimate-3-phosphate synthase into chloroplasts of higher plants *in vitro*. *Proc. Natl. Acad. Sci. USA* **83**, 6873-6877.
- Delwiche, C.F., Kuhsel, M., and Palmer, J.D.** (1995). Phylogenetic analysis of tufA sequences indicates a cyanobacterial origin of all plastids. *Mol Phylogenet Evol* **4**, 110-128.
- Derr, L.K., Strathern, J.N., and Garfinkel, D.J.** (1991). RNA-mediated recombination in *S. cerevisiae*. *Cell* **67**, 355-364.
- Douce, R., and Joyard, J.** (1990). Biochemistry and function of the plastid envelope. *Annu Rev Cell Biol* **6**, 173-216.
- Drouin, G., and Moniz de Sa, M.** (1997). Loss of introns in the pollen-specific actin subfamily members of potato and tomato. *J. Mol. Evol.* **45**, 509-513.
- Duchêne, A.M., Giritch, A., Hoffmann, B., Cognat, V., Lancelin, D., Peeters, N.M., Zaepfel, M., Marechal-Drouard, L., and Small, I.D.** (2005). Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **102**, 16484-16489.
- Duy, D., Wanner, G., Meda, A.R., von Wiren, N., Soll, J., and Philippar, K.** (2007). PIC1, an ancient permease in *Arabidopsis* chloroplasts, mediates iron transport. *Plant Cell* **19**, 986-1006.
- Edwards, K., Johnstone, C., and Thompson, C.** (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res.* **19**, 1349.
- Emanuelsson, O., Nielsen, H., and von Heijne, G.** (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* **8**, 978-984.
- Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G.** (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* **300**, 1005-1016.
- Errington, J., Daniel, R.A., and Scheffers, D.J.** (2003). Cytokinesis in bacteria. *Microbiol Mol Biol Rev* **67**, 52-65, table of contents.
- Ertel, F., Mirus, O., Bredemeier, R., Moslavac, S., Becker, T., and Schleiff, E.** (2005). The evolutionarily related beta-barrel polypeptide transporters from *Pisum sativum* and *Nostoc PCC7120* contain two distinct functional domains. *J. Biol. Chem.* **280**, 28281-28289.
- Flügge, U.I., and Hinz, G.** (1986). Energy dependence of protein translocation into chloroplasts. *Eur J Biochem* **160**, 563-570.
- Friedman, A.L., and Keegstra, K.** (1989). Chloroplast protein import: quantitative analysis of precursor binding. *Plant Physiol.* **89**, 993-999.

- Gentle, I.E., Burri, L., and Lithgow, T.** (2005). Molecular architecture and function of the Omp85 family of proteins. *Mol. Microbiol.* **58**, 1216-1225.
- Golenberg, E.M., Clegg, M.T., Durbin, M.L., Doebley, J., and Ma, D.P.** (1993). Evolution of a noncoding region of the chloroplast genome. *Mol Phylogenet Evol* **2**, 52-64.
- Gray, M.W.** (1999). Evolution of organellar genomes. *Curr Opin Genet Dev* **9**, 678-687.
- Grennan, A.K.** (2006). Genevestigator. Facilitating web-based gene-expression analysis. *Plant Physiol.* **141**, 1164-1166.
- Gutensohn, M., Schulz, B., Nicolay, P., and Flügge, U.I.** (2000). Functional analysis of the two *Arabidopsis* homologues of Toc34, a component of the chloroplast protein import apparatus. *Plant J.* **23**, 771-783.
- Häger, K.P., Müller, B., Wind, C., Erbach, S., and Fischer, H.** (1996). Evolution of legumin genes: loss of an ancestral intron at the beginning of angiosperm diversification. *FEBS Lett.* **387**, 94-98.
- Halperin, T., Ostersetzer, O., and Adam, Z.** (2001). ATP-dependent association between subunits of Clp protease in pea chloroplasts. *Planta* **213**, 614-619.
- Hartel, H., Lokstein, H., Dormann, P., Grimm, B., and Benning, C.** (1997). Changes in the composition of the photosynthetic apparatus in the galactolipid-deficient *dgd1* mutant of *Arabidopsis thaliana*. *Plant Physiol.* **115**, 1175-1184.
- Hartl, F.U., Schmidt, B., Wachter, E., Weiss, H., and Neupert, W.** (1986). Transport into mitochondria and intramitochondrial sorting of the Fe/S protein of ubiquinol-cytochrome c reductase. *Cell* **47**, 939-951.
- Heins, L., Mehrle, A., Hemmler, R., Wagner, R., Kuchler, M., Hörmann, F., Sveshnikov, D., and Soll, J.** (2002). The preprotein conducting channel at the inner envelope membrane of plastids. *EMBO J.* **21**, 2616-2625.
- Herrmann, J.M., and Neupert, W.** (2000). Protein transport into mitochondria. *Curr Opin Microbiol* **3**, 210-214.
- Hill, K., Model, K., Ryan, M.T., Dietmeier, K., Martin, F., Wagner, R., and Pfanner, N.** (1998). Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins. *Nature* **395**, 516-521.
- Hiltbrunner, A., Bauer, J., Vidi, P.A., Infanger, S., Weibel, P., Hohwy, M., and Kessler, F.** (2001). Targeting of an abundant cytosolic form of the protein import receptor at Toc159 to the outer chloroplast membrane. *J. Cell Biol.* **154**, 309-316.
- Hinnah, S.C., Hill, K., Wagner, R., Schlicher, T., and Soll, J.** (1997). Reconstitution of a chloroplast protein import channel. *EMBO J.* **16**, 7351-7360.
- Hinnah, S.C., Wagner, R., Sveshnikova, N., Harrer, R., and Soll, J.** (2002). The chloroplast protein import channel Toc75: pore properties and interaction with transit peptides. *Biophys. J.* **83**, 899-911.
- Hirsch, S., Muckel, E., Heemeyer, F., von Heijne, G., and Soll, J.** (1994). A receptor component of the chloroplast protein translocation machinery. *Science* **266**, 1989-1992.
- Hofmann, N.R., and Theg, S.M.** (2005a). Protein- and energy-mediated targeting of chloroplast outer envelope membrane proteins. *Plant J.* **44**, 917-927.
- Hofmann, N.R., and Theg, S.M.** (2005b). Chloroplast outer membrane protein targeting and insertion. *Trends Plant Sci.* **10**, 450-457.
- Hofmann, N.R., and Theg, S.M.** (2005c). Toc64 is not required for import of proteins into chloroplasts in the moss *Physcomitrella patens*. *Plant J* **43**, 675-687.

- Hörmann, F., Küchler, M., Sveshnikov, D., Oppermann, U., Li, Y., and Soll, J.** (2004). Tic32, an essential component in chloroplast biogenesis. *J. Biol. Chem.* **279**, 34756-34762.
- Horniak, L., Pilon, M., van 't Hof, R., and de Kruijff, B.** (1993). The secondary structure of the ferredoxin transit sequence is modulated by its interaction with negatively charged lipids. *FEBS Lett* **334**, 241-246.
- Horst, M., Hilfiker-Rothenfluh, S., Oppliger, W., and Schatz, G.** (1995). Dynamic interaction of the protein translocation systems in the inner and outer membranes of yeast mitochondria. *EMBO J.* **14**, 2293-2297.
- Howden, R., Park, S.K., Moore, J.M., Orme, J., Grossniklaus, U., and Twell, D.** (1998). Selection of T-DNA-tagged male and female gametophytic mutants by segregation distortion in *Arabidopsis*. *Genetics* **149**, 621-631.
- Huang, N., Sutliff, T.D., Litts, J.C., and Rodriguez, R.L.** (1990). Classification and characterization of the rice alpha-amylase multigene family. *Plant Mol Biol* **14**, 655-668.
- Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O'Shea, E.K.** (2003). Global analysis of protein localization in budding yeast. *Nature* **425**, 686-691.
- Hust, B., and Gutensohn, M.** (2006). Deletion of core components of the plastid protein import machinery causes differential arrest of embryo development in *Arabidopsis thaliana*. *Plant Biol. (Stuttg.)* **8**, 18-30.
- Inaba, T., and Schnell, D.J.** (2008). Protein trafficking to plastids: one theme, many variations. *Biochem J.* **413**, 15-28.
- Inaba, T., Li, M., Alvarez-Huerta, M., Kessler, F., and Schnell, D.J.** (2003). atTic110 functions as a scaffold for coordinating the stromal events of protein import into chloroplasts. *J. Biol. Chem.* **278**, 38617-38627.
- Inaba, T., Alvarez-Huerta, M., Li, M., Bauer, J., Ewers, C., Kessler, F., and Schnell, D.J.** (2005). *Arabidopsis* Tic110 is essential for the assembly and function of the protein import machinery of plastids. *Plant Cell* **17**, 1482-1496.
- Inoue, H., Nojima, H., and Okayama, H.** (1990). High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**, 23-28.
- Inoue, K.** (2007). The chloroplast outer envelope membrane: the edge of light and excitement. *J. Integr. Plant Biol.* **49**, 1100-1111.
- Inoue, K., and Potter, D.** (2004). The chloroplastic protein translocation channel Toc75 and its paralog OEP80 represent two distinct protein families and are targeted to the chloroplastic outer envelope by different mechanisms. *Plant J.* **39**, 354-365.
- Inoue, K., Demel, R., de Kruijff, B., and Keegstra, K.** (2001). The N-terminal portion of the preToc75 transit peptide interacts with membrane lipids and inhibits binding and import of precursor proteins into isolated chloroplasts. *Eur. J. Biochem.* **268**, 4036-4043.
- Ivey, R.A., 3rd, Subramanian, C., and Bruce, B.D.** (2000). Identification of a Hsp70 recognition domain within the rubisco small subunit transit peptide. *Plant Physiol* **122**, 1289-1299.
- Jackson, D.T., Froehlich, J.E., and Keegstra, K.** (1998). The hydrophilic domain of Tic110, an inner envelope membrane component of the chloroplastic protein translocation apparatus, faces the stromal compartment. *J. Biol. Chem.* **273**, 16583-16588.

- Jackson-Constan, D., and Keegstra, K.** (2001). Arabidopsis genes encoding components of the chloroplastic protein import apparatus. *Plant Physiol.* **125**, 1567-1576.
- Jackson-Constan, D., Akita, M., and Keegstra, K.** (2001). Molecular chaperones involved in chloroplast protein import. *Biochim. Biophys. Acta* **1541**, 102-113.
- Jarvis, P.** (2004). Organellar proteomics: chloroplasts in the spotlight. *Curr. Biol.* **14**, R317-R319.
- Jarvis, P.** (2008a). Targeting of nucleus-encoded proteins to chloroplasts in plants. *New Phytol.* **179**, 257-285.
- Jarvis, P.** (2008b). Targeting of nucleus-encoded proteins to chloroplasts in plants (Tansley Review). *New Phytol.* **179**, 257-285.
- Jarvis, P., and Soll, J.** (2001). Toc, Tic, and chloroplast protein import. *Biochim. Biophys. Acta* **1541**, 64-79.
- Jarvis, P., Chen, L.J., Li, H., Peto, C.A., Fankhauser, C., and Chory, J.** (1998). An *Arabidopsis* mutant defective in the plastid general protein import apparatus. *Science* **282**, 100-103.
- Jelic, M., Soll, J., and Schleiff, E.** (2003). Two Toc34 homologues with different properties. *Biochemistry* **42**, 5906-5916.
- Jelic, M., Sveshnikova, N., Motzkus, M., Horth, P., Soll, J., and Schleiff, E.** (2002). The chloroplast import receptor Toc34 functions as preprotein-regulated GTPase. *Biol. Chem.* **383**, 1875-1883.
- Jensen, R.E., and Johnson, A.E.** (1999). Protein translocation: is Hsp70 pulling my chain? *Curr Biol* **9**, R779-782.
- Joyard, J., Billecocq, A., Bartlett, S.G., Block, M.A., Chua, N.H., and Douce, R.** (1983). Localization of polypeptides to the cytosolic side of the outer envelope membrane of spinach chloroplasts. *J Biol Chem* **258**, 10000-10006.
- Joyard, J., Teyssier, E., Miege, C., Berny-Seigneurin, D., Marechal, E., Block, M.A., Dorne, A.J., Rolland, N., Ajlani, G., and Douce, R.** (1998). The biochemical machinery of plastid envelope membranes. *Plant Physiol* **118**, 715-723.
- Kalanon, M., and McFadden, G.I.** (2008). The chloroplast protein translocation complexes of *Chlamydomonas reinhardtii*: a bioinformatic comparison of Toc and Tic components in plants, green algae and red algae. *Genetics* **179**, 95-112.
- Kalanon, M., Tonkin, C.J., and McFadden, G.I.** (2009). Characterization of two putative protein translocation components in the apicoplast of *Plasmodium falciparum*. *Eukaryot Cell* **8**, 1146-1154.
- Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., and Tabata, S.** (1996). Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* **3**, 109-136.
- Karimi, M., De Meyer, B., and Hilson, P.** (2005). Modular cloning in plant cells. *Trends Plant Sci.* **10**, 103-105.
- Keegstra, K.** (1989). Transport and routing of proteins into chloroplasts. *Cell* **56**, 247-253.
- Keeling, P.** (2004). A brief history of plastids and their hosts. *Protist* **155**, 3-7.

- Kessler, F., and Blobel, G.** (1996). Interaction of the protein import and folding machineries of the chloroplast. *Proc. Natl. Acad. Sci. USA* **93**, 7684-7689.
- Kessler, F., and Schnell, D.J.** (2006a). The function and diversity of plastid protein import pathways: a multilane GTPase highway into plastids. *Traffic* **7**, 248-257.
- Kessler, F., and Schnell, D.J.** (2006b). The function and diversity of plastid protein import pathways: a multilane GTPase highway into plastids. *Traffic* **7**, 248-257.
- Kessler, F., Blobel, G., Patel, H.A., and Schnell, D.J.** (1994). Identification of two GTP-binding proteins in the chloroplast protein import machinery. *Science* **266**, 1035-1039.
- Kikuchi, S., Hirohashi, T., and Nakai, M.** (2006). Characterization of the preprotein translocon at the outer envelope membrane of chloroplasts by blue native PAGE. *Plant Cell Physiol* **47**, 363-371.
- Kikuchi, S., Oishi, M., Hirabayashi, Y., Lee, D.W., Hwang, I., and Nakai, M.** (2009). A 1-megadalton translocation complex containing Tic20 and Tic21 mediates chloroplast protein import at the inner envelope membrane. *Plant Cell* **21**, 1781-1797.
- Kirk, J.T.O., and Tilney-Bassett, R.A.E.** (1978). The plastids: their chemistry, structure, growth, and inheritance. Elsevier/North-Holland Biomedical Press **2**.
- Ko, K., Budd, D., Wu, C., Seibert, F., Kourtz, L., and Ko, Z.W.** (1995). Isolation and characterization of a cDNA clone encoding a member of the Com44/Cim44 envelope components of the chloroplast protein import apparatus. *J. Biol. Chem.* **270**, 28601-28608.
- Kouranov, A., and Schnell, D.J.** (1997). Analysis of the interactions of preproteins with the import machinery over the course of protein import into chloroplasts. *J. Cell Biol.* **139**, 1677-1685.
- Kouranov, A., Wang, H., and Schnell, D.J.** (1999). Tic22 is targeted to the intermembrane space of chloroplasts by a novel pathway. *J. Biol. Chem.* **274**, 25181-25186.
- Kouranov, A., Chen, X., Fuks, B., and Schnell, D.J.** (1998). Tic20 and Tic22 are new components of the protein import apparatus at the chloroplast inner envelope membrane. *J. Cell Biol.* **143**, 991-1002.
- Kourtz, L., and Ko, K.** (1997). The early stage of chloroplast protein import involves Com70. *J Biol Chem* **272**, 2808-2813.
- Kovacheva, S., Bedard, J., Wardle, A., Patel, R., and Jarvis, P.** (2007). Further in vivo studies on the role of the molecular chaperone, Hsp93, in plastid protein import. *Plant J* **50**, 364-379.
- Kovacheva, S., Bédard, J., Patel, R., Dudley, P., Twell, D., Ríos, G., Koncz, C., and Jarvis, P.** (2005). *In vivo* studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. *Plant J.* **41**, 412-428.
- Kovermann, P., Truscott, K.N., Guiard, B., Rehling, P., Sepuri, N.B., Muller, H., Jensen, R.E., Wagner, R., and Pfanner, N.** (2002). Tim22, the essential core of the mitochondrial protein insertion complex, forms a voltage-activated and signal-gated channel. *Mol. Cell* **9**, 363-373.
- Krimm, I., Gans, P., Hernandez, J.F., Arlaud, G.J., and Lancelin, J.M.** (1999). A coil-helix instead of a helix-coil motif can be induced in a chloroplast transit peptide from *Chlamydomonas reinhardtii*. *Eur. J. Biochem.* **265**, 171-180.
- Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L.** (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567-580.

- Kubis, S., Baldwin, A., Patel, R., Razzaq, A., Dupree, P., Lilley, K., Kurth, J., Leister, D., and Jarvis, P.** (2003). The *Arabidopsis ppil* mutant is specifically defective in the expression, chloroplast import, and accumulation of photosynthetic proteins. *Plant Cell* **15**, 1859-1871.
- Kubis, S., Patel, R., Combe, J., Bédard, J., Kovacheva, S., Lilley, K., Biehl, A., Leister, D., Ríos, G., Koncz, C., and Jarvis, P.** (2004). Functional specialization amongst the *Arabidopsis* Toc159 family of chloroplast protein import receptors. *Plant Cell* **16**, 2059-2077.
- Kubis, S.E., Lilley, K.S., and Jarvis, P.** (2008). Isolation and preparation of chloroplasts from *Arabidopsis thaliana* plants. *Methods Mol Biol* **425**, 171-186.
- Küchler, M., Decker, S., Hörmann, F., Soll, J., and Heins, L.** (2002). Protein import into chloroplasts involves redox-regulated proteins. *EMBO J.* **21**, 6136-6145.
- Kumar, A., Agarwal, S., Heyman, J.A., Matson, S., Heidtman, M., Piccirillo, S., Umansky, L., Drawid, A., Jansen, R., Liu, Y., Cheung, K.H., Miller, P., Gerstein, M., Roeder, G.S., and Snyder, M.** (2002). Subcellular localization of the yeast proteome. *Genes Dev* **16**, 707-719.
- Kumar, V., and Trick, M.** (1993). Sequence complexity of the S receptor kinase gene family in *Brassica*. *Mol. Gen. Genet.* **241**, 440-446.
- Laemmli, U.K.** (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.
- Lee, Y.J., Sohn, E.J., Lee, K.H., Lee, D.W., and Hwang, I.** (2004). The transmembrane domain of AtToc64 and its C-terminal lysine-rich flanking region are targeting signals to the chloroplast outer envelope membrane. *Mol. Cells* **17**, 281-291.
- Leister, D.** (2003). Chloroplast research in the genomic age. *Trends Genet.* **19**, 47-56.
- Lesk, A.M., and Fordham, W.D.** (1996). Conservation and variability in the structures of serine proteinases of the chymotrypsin family. *J. Mol. Biol.* **258**, 501-537.
- Li, H.M., Kesavulu, M.M., Su, P.H., Yeh, Y.H., and Hsiao, C.D.** (2007). Toc GTPases. *J. Biomed. Sci.* **14**, 505-508.
- Lopez-Juez, E., and Pyke, K.A.** (2005). Plastids unleashed: their development and their integration in plant development. *Int. J. Dev. Biol.* **49**, 557-577.
- Lubben, T.H., Donaldson, G.K., Viitanen, P.V., and Gatenby, A.A.** (1989). Several proteins imported into chloroplasts form stable complexes with the GroEL-related chloroplast molecular chaperone. *Plant Cell* **1**, 1223-1230.
- Lübeck, J., Soll, J., Akita, M., Nielsen, E., and Keegstra, K.** (1996). Topology of IEP110, a component of the chloroplastic protein import machinery present in the inner envelope membrane. *EMBO J.* **15**, 4230-4238.
- Ma, Y., Kouranov, A., LaSala, S.E., and Schnell, D.J.** (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.
- Marshall, J.S., DeRocher, A.E., Keegstra, K., and Vierling, E.** (1990). Identification of heat shock protein hsp70 homologues in chloroplasts. *Proc. Natl. Acad. Sci. USA* **87**, 374-378.
- Martin, J., Mahlke, K., and Pfanner, N.** (1991). Role of an energized inner membrane in mitochondrial protein import. Delta psi drives the movement of presequences. *J. Biol. Chem.* **266**, 18051-18057.
- Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M., and Penny, D.** (2002). Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny

- and thousands of cyanobacterial genes in the nucleus. *Proc. Natl. Acad. Sci. USA* **99**, 12246-12251.
- Mason, J.R., and Cammack, R.** (1992). The electron-transport proteins of hydroxylating bacterial dioxygenases. *Annu Rev Microbiol* **46**, 277-305.
- Mathews, D.H., Sabina, J., Zuker, M., and Turner, D.H.** (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**, 911-940.
- May, T., and Soll, J.** (2000). 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. *Plant Cell* **12**, 53-64.
- McCormac, A.C., Elliott, M.C., and Chen, D.F.** (1998). A simple method for the production of highly competent cells of *Agrobacterium* for transformation via electroporation. *Mol. Biotechnol.* **9**, 155-159.
- McFadden, G.I.** (1999). Endosymbiosis and evolution of the plant cell. *Curr. Opin. Plant Biol.* **2**, 513-519.
- Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K., Marechal-Drouard, L., Marshall, W.F., Qu, L.H., Nelson, D.R., Sanderfoot, A.A., Spalding, M.H., Kapitonov, V.V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S.M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C.L., Cognat, V., Croft, M.T., Dent, R., Dutcher, S., Fernandez, E., Fukuzawa, H., Gonzalez-Ballester, D., Gonzalez-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P.A., Lemaire, S.D., Lobanov, A.V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J.V., Moseley, J., Napoli, C., Nedelcu, A.M., Niyogi, K., Novoselov, S.V., Paulsen, I.T., Pazour, G., Purton, S., Ral, J.P., Riano-Pachon, D.M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S.L., Allmer, J., Balk, J., Bisova, K., Chen, C.J., Elias, M., Gendler, K., Hauser, C., Lamb, M.R., Ledford, H., Long, J.C., Minagawa, J., Page, M.D., Pan, J., Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A.M., Yang, P., Ball, S., Bowler, C., Dieckmann, C.L., Gladyshev, V.N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R.T., Brokstein, P., *et al.* (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* **318**, 245-250.**
- Miège, C., Marechal, E., Shimojima, M., Awai, K., Block, M.A., Ohta, H., Takamiya, K., Douce, R., and Joyard, J.** (1999). Biochemical and topological properties of type A MGDG synthase, a spinach chloroplast envelope enzyme catalyzing the synthesis of both prokaryotic and eukaryotic MGDG. *Eur J Biochem* **265**, 990-1001.
- Moberg, P., Stahl, A., Bhushan, S., Wright, S.J., Eriksson, A., Bruce, B.D., and Glaser, E.** (2003). Characterization of a novel zinc metalloprotease involved in degrading targeting peptides in mitochondria and chloroplasts. *Plant J.* **36**, 616-628.
- Morton, B.R.** (1999). Strand asymmetry and codon usage bias in the chloroplast genome of *Euglena gracilis*. *Proc Natl Acad Sci U S A* **96**, 5123-5128.
- Murashige, T., and Skoog, F.** (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant* **15**, 473-497.
- Nakrieko, K.A., Mould, R.M., and Smith, A.G.** (2004). Fidelity of targeting to chloroplasts is not affected by removal of the phosphorylation site from the transit peptide. *Eur. J. Biochem.* **271**, 509-516.

- Nelson, N., and Ben-Shem, A.** (2004). The complex architecture of oxygenic photosynthesis. *Nat. Rev. Mol. Cell Biol.* **5**, 971-982.
- Neuhaus, H.E., and Emes, M.J.** (2000). Nonphotosynthetic metabolism in plastids. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **51**, 111-140.
- Neupert, W., and Brunner, M.** (2002). The protein import motor of mitochondria. *Nat. Rev. Mol. Cell Biol.* **3**, 555-565.
- Neupert, W., and Herrmann, J.M.** (2007). Translocation of proteins into mitochondria. *Annu. Rev. Biochem.* **76**, 723-749.
- Nielsen, E., Akita, M., Davila-Aponte, J., and Keegstra, K.** (1997). Stable association of chloroplastic precursors with protein translocation complexes that contain proteins from both envelope membranes and a stromal Hsp100 molecular chaperone. *EMBO J.* **16**, 935-946.
- Niewiadomski, P., Knappe, S., Geimer, S., Fischer, K., Schulz, B., Unte, U.S., Rosso, M.G., Ache, P., Flügge, U.I., and Schneider, A.** (2005). The Arabidopsis plastidic glucose 6-phosphate/phosphate translocator GPT1 is essential for pollen maturation and embryo sac development. *Plant Cell* **17**, 760-775.
- Olsen, L.J., and Keegstra, K.** (1992). The binding of precursor proteins to chloroplasts requires nucleoside triphosphates in the intermembrane space. *J. Biol. Chem.* **267**, 433-439.
- Pain, D., and Blobel, G.** (1987). Protein import into chloroplasts requires a chloroplast ATPase. *Proc. Natl. Acad. Sci. USA* **84**, 3288-3292.
- Pan, S., Carter, C.J., and Raikhel, N.V.** (2005). Understanding protein trafficking in plant cells through proteomics. *Expert Rev. Proteomics* **2**, 781-792.
- Patel, R., Hsu, S., Bédard, J., Inoue, K., and Jarvis, P.** (2008). The Omp85-related chloroplast outer envelope protein OEP80 is essential for viability in Arabidopsis. *Plant Physiol.* **148**, 235-245.
- Perry, S.E., and Keegstra, K.** (1994). Envelope membrane proteins that interact with chloroplastic precursor proteins. *Plant Cell* **6**, 93-105.
- Pilon, M., and Schekman, R.** (1999). Protein translocation: how Hsp70 pulls it off. *Cell* **97**, 679-682.
- Pinnaduwege, P., and Bruce, B.D.** (1996). In vitro interaction between a chloroplast transit peptide and chloroplast outer envelope lipids is sequence-specific and lipid class- dependent. *J Biol Chem* **271**, 32907-32915.
- Pollmann, S., Neu, D., Lehmann, T., Berkowitz, O., Schafer, T., and Weiler, E.W.** (2006). Subcellular localization and tissue specific expression of amidase 1 from *Arabidopsis thaliana*. *Planta* **224**, 1241-1253.
- Porra, R.J., Thompson, W.A., and Kriedemann, P.E.** (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls *a* and *b* extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochim. Biophys. Acta* **975**, 384-394.
- Puchta, H.** (2002). Gene replacement by homologous recombination in plants. *Plant Mol Biol* **48**, 173-182.
- Qbadou, S., Tien, R., Soll, J., and Schleiff, E.** (2003). Membrane insertion of the chloroplast outer envelope protein, Toc34: constrains for insertion and topology. *J Cell Sci* **116**, 837-846.
- Qbadou, S., Becker, T., Mirus, O., Tews, I., Soll, J., and Schleiff, E.** (2006). The molecular chaperone Hsp90 delivers precursor proteins to the chloroplast import receptor Toc64. *EMBO J.* **25**, 1836-1847.

- Qbadou, S., Becker, T., Bionda, T., Reger, K., Ruprecht, M., Soll, J., and Schleiff, E.** (2007). Toc64 - a preprotein-receptor at the outer membrane with bipartite function. *J. Mol. Biol.* **367**, 1330-1346.
- Rapaport, D., and Neupert, W.** (1999). Biogenesis of Tom40, core component of the TOM complex of mitochondria. *J Cell Biol* **146**, 321-331.
- Rassow, J., Dekker, P.J., van Wilpe, S., Meijer, M., and Soll, J.** (1999). The preprotein translocase of the mitochondrial inner membrane: function and evolution. *J. Mol. Biol.* **286**, 105-120.
- Reumann, S., and Keegstra, K.** (1999). The endosymbiotic origin of the protein import machinery of chloroplastic envelope membranes. *Trends Plant Sci.* **4**, 302-307.
- Reumann, S., Davila-Aponte, J., and Keegstra, K.** (1999). The evolutionary origin of the protein-translocating channel of chloroplastic envelope membranes: identification of a cyanobacterial homolog. *Proc. Natl. Acad. Sci. USA* **96**, 784-789.
- Reumann, S., Inoue, K., and Keegstra, K.** (2005). Evolution of the general protein import pathway of plastids. *Mol. Membr. Biol.* **22**, 73-86.
- Rial, D.V., Arakaki, A.K., and Ceccarelli, E.A.** (2000). Interaction of the targeting sequence of chloroplast precursors with Hsp70 molecular chaperones. *Eur. J. Biochem.* **267**, 6239-6248.
- Rial, D.V., Ottado, J., and Ceccarelli, E.A.** (2003). Precursors with altered affinity for Hsp70 in their transit peptides are efficiently imported into chloroplasts. *J. Biol. Chem.* **278**, 46473-46481.
- Rial, D.V., Lombardo, V.A., Ceccarelli, E.A., and Ottado, J.** (2002). The import of ferredoxin-NADP⁺ reductase precursor into chloroplasts is modulated by the region between the transit peptide and the mature core of the protein. *Eur J Biochem* **269**, 5431-5439.
- Richter, S., and Lamppa, G.K.** (1999). Stromal processing peptidase binds transit peptides and initiates their ATP-dependent turnover in chloroplasts. *J. Cell Biol.* **147**, 33-44.
- Richter, S., and Lamppa, G.K.** (2002). Determinants for removal and degradation of transit peptides of chloroplast precursor proteins. *J. Biol. Chem.* **277**, 43888-43894.
- Richter, S., and Lamppa, G.K.** (2003). Structural properties of the chloroplast stromal processing peptidase required for its function in transit peptide removal. *J. Biol. Chem.* **278**, 39497-39502.
- Roth, R.A.** (2004). *Handbook of proteolytic enzymes*, **2 edn**, 868 - 871.
- Rothfield, L.I., and Zhao, C.R.** (1996). How do bacteria decide where to divide? *Cell* **84**, 183-186.
- Rudhe, C., Clifton, R., Chew, O., Zemam, K., Richter, S., Lamppa, G., Whelan, J., and Glaser, E.** (2004). Processing of the dual targeted precursor protein of glutathione reductase in mitochondria and chloroplasts. *J. Mol. Biol.* **343**, 639-647.
- Sánchez-Pulido, L., Devos, D., Genevrois, S., Vicente, M., and Valencia, A.** (2003). POTRA: a conserved domain in the FtsQ family and a class of beta-barrel outer membrane proteins. *Trends Biochem. Sci.* **28**, 523-526.
- Schatz, G., and Dobberstein, B.** (1996). Common principles of protein translocation across membranes. *Science* **271**, 1519-1526.

- Schirmer, E.C., Glover, J.R., Singer, M.A., and Lindquist, S.** (1996). HSP100/Clp proteins: a common mechanism explains diverse functions. *Trends Biochem. Sci.* **21**, 289-296.
- Schleiff, E., and Klösgen, R.B.** (2001). Without a little help from 'my' friends: direct insertion of proteins into chloroplast membranes? *Biochim. Biophys. Acta* **1541**, 22-33.
- Schleiff, E., Jelic, M., and Soll, J.** (2003a). A GTP-driven motor moves proteins across the outer envelope of chloroplasts. *Proc. Natl. Acad. Sci. USA* **100**, 4604-4609.
- Schleiff, E., Soll, J., Kuchler, M., Kuhlbrandt, W., and Harrer, R.** (2003b). Characterization of the translocon of the outer envelope of chloroplasts. *J. Cell Biol.* **160**, 541-551.
- Schnell, D.J., and Blobel, G.** (1993). Identification of intermediates in the pathway of protein import into chloroplasts and their localization to envelope contact sites. *J. Cell Biol.* **120**, 103-115.
- Schnell, D.J., and Hebert, D.N.** (2003). Protein translocons: multifunctional mediators of protein translocation across membranes. *Cell* **112**, 491-505.
- Schnell, D.J., Kessler, F., and Blobel, G.** (1994). Isolation of components of the chloroplast protein import machinery. *Science* **266**, 1007-1012.
- Schnell, D.J., Blobel, G., Keegstra, K., Kessler, F., Ko, K., and Soll, J.** (1997). A consensus nomenclature for the protein-import components of the chloroplast envelope. *Trends Cell Biol* **7**, 303-304.
- Schröder, W.P., and Kieselbach, T.** (2003). Update on chloroplast proteomics. *Photosynth Res* **78**, 181-193.
- Seedorf, M., Waegemann, K., and Soll, J.** (1995). A constituent of the chloroplast import complex represents a new type of GTP-binding protein. *Plant J.* **7**, 401-411.
- Shanklin, J., DeWitt, N.D., and Flanagan, J.M.** (1995). The stroma of higher plant plastids contain ClpP and ClpC, functional homologs of *Escherichia coli* ClpP and ClpA: an archetypal two-component ATP-dependent protease. *Plant Cell* **7**, 1713-1722.
- Silva-Filho, M.C.** (2003). One ticket for multiple destinations: dual targeting of proteins to distinct subcellular locations. *Curr. Opin. Plant Biol.* **6**, 589-595.
- Simon, A., Romert, A., Gustafson, A.L., McCaffery, J.M., and Eriksson, U.** (1999). Intracellular localization and membrane topology of 11-cis retinol dehydrogenase in the retinal pigment epithelium suggest a compartmentalized synthesis of 11-cis retinaldehyde. *J. Cell Sci.* **112**, 549-558.
- Smith, M.D.** (2006). Protein import into chloroplasts: an ever-evolving story. *Can. J. Bot.* **84**, 531-542.
- Smith, M.D., Rounds, C.M., Wang, F., Chen, K., Afithile, M., and Schnell, D.J.** (2004). atToc159 is a selective transit peptide receptor for the import of nucleus-encoded chloroplast proteins. *J. Cell Biol.* **165**, 323-334.
- Sohrt, K., and Soll, J.** (2000). Toc64, a new component of the protein translocon of chloroplasts. *J. Cell Biol.* **148**, 1213-1221.
- Sokolenko, A., Lerbs-Mache, S., Altschmied, L., and Herrmann, R.G.** (1998). Clp protease complexes and their diversity in chloroplasts. *Planta* **207**, 286-295.
- Soll, J., and Schleiff, E.** (2004). Protein import into chloroplasts. *Nat. Rev. Mol. Cell Biol.* **5**, 198-208.
- Somerville, C., and Koornneef, M.** (2002). A fortunate choice: the history of Arabidopsis as a model plant. *Nat. Rev. Genet.* **3**, 883-889.

- Stahl, T., Glockmann, C., Soll, J., and Heins, L.** (1999). Tic40, a new "old" subunit of the chloroplast protein import translocon. *J. Biol. Chem.* **274**, 37467-37472.
- Stengel, A., Benz, P., Balsera, M., Soll, J., and Bolter, B.** (2008). TIC62 redox-regulated translocon composition and dynamics. *J Biol Chem* **283**, 6656-6667.
- Strub, A., Lim, J.H., Pfanner, N., and Voos, W.** (2000). The mitochondrial protein import motor. *Biol Chem* **381**, 943-949.
- Subramanian, C., Ivey, R., 3rd, and Bruce, B.D.** (2001). Cytometric analysis of an epitope-tagged transit peptide bound to the chloroplast translocation apparatus. *Plant J.* **25**, 349-363.
- Sun, Y.J., Forouhar, F., Li Hm, H.M., Tu, S.L., Yeh, Y.H., Kao, S., Shr, H.L., Chou, C.C., Chen, C., and Hsiao, C.D.** (2002). Crystal structure of pea Toc34, a novel GTPase of the chloroplast protein translocon. *Nat. Struct. Biol.* **9**, 95-100.
- Sveshnikova, N., Soll, J., and Schleiff, E.** (2000a). Toc34 is a preprotein receptor regulated by GTP and phosphorylation. *Proc. Natl. Acad. Sci. USA* **97**, 4973-4978.
- Sveshnikova, N., Grimm, R., Soll, J., and Schleiff, E.** (2000b). Topology studies of the chloroplast protein import channel Toc75. *Biol Chem* **381**, 687-693.
- Teng, Y.S., Su, Y.S., Chen, L.J., Lee, Y.J., Hwang, I., and Li, H.M.** (2006). Tic21 is an essential translocon component for protein translocation across the chloroplast inner envelope membrane. *Plant Cell* **18**, 2247-2257.
- The-Arabidopsis-Genome-Initiative.** (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796-815.
- Theg, S.M., Bauerle, C., Olsen, L.J., Selman, B.R., and Keegstra, K.** (1989). Internal ATP is the only energy requirement for the translocation of precursor proteins across chloroplastic membranes. *J. Biol. Chem.* **264**, 6730-6736.
- Tranel, P.J., and Keegstra, K.** (1996). A novel, bipartite transit peptide targets OEP75 to the outer membrane of the chloroplastic envelope. *Plant Cell* **8**, 2093-2104.
- Tranel, P.J., Froehlich, J., Goyal, A., and Keegstra, K.** (1995). A component of the chloroplastic protein import apparatus is targeted to the outer envelope membrane via a novel pathway. *EMBO J.* **14**, 2436-2446.
- Truscott, K.N., Kovermann, P., Geissler, A., Merlin, A., Meijer, M., Driessen, A.J., Rassow, J., Pfanner, N., and Wagner, R.** (2001). A presequence- and voltage-sensitive channel of the mitochondrial preprotein translocase formed by Tim23. *Nat. Struct. Biol.* **8**, 1074-1082.
- Tsugeki, R., and Nishimura, M.** (1993). Interaction of homologues of Hsp70 and Cpn60 with ferredoxin-NADP⁺ reductase upon its import into chloroplasts. *FEBS Lett* **320**, 198-202.
- van Dooren, G.G., Tomova, C., Agrawal, S., Humbel, B.M., and Striepen, B.** (2008). *Toxoplasma gondii* Tic20 is essential for apicoplast protein import. *Proc Natl Acad Sci U S A* **105**, 13574-13579.
- van Wijk, K.J.** (2004). Plastid proteomics. *Plant Physiol. Biochem.* **42**, 963-977.
- VanderVere, P.S., Bennett, T.M., Oblong, J.E., and Lamppa, G.K.** (1995). A chloroplast processing enzyme involved in precursor maturation shares a zinc-binding motif with a recently recognized family of metalloendopeptidases. *Proc. Natl. Acad. Sci. USA* **92**, 7177-7181.
- van't Hof, R., van Klompenburg, W., Pilon, M., Kozubek, A., de Korte-Kool, G., Demel, R.A., Weisbeek, P.J., and de Kruijff, B.** (1993). The transit sequence mediates the specific interaction of the precursor of ferredoxin with chloroplast envelope membrane lipids. *J Biol Chem* **268**, 4037-4042.

- von Heijne, G., and Nishikawa, K.** (1991). Chloroplast transit peptides. The perfect random coil? *FEBS Lett.* **278**, 1-3.
- Wallas, T.R., Smith, M.D., Sanchez-Nieto, S., and Schnell, D.J.** (2003). The roles of *toc34* and *toc75* in targeting the *toc159* preprotein receptor to chloroplasts. *J. Biol. Chem.* **278**, 44289-44297.
- Wan, J., Bringloe, D., and Lamppa, G.K.** (1998). Disruption of chloroplast biogenesis and plant development upon down-regulation of a chloroplast processing enzyme involved in the import pathway. *Plant J.* **15**, 459-468.
- Whatley, J.M.** (1978). A suggested cycle of plastid developmental interrelationships. *New Phytol.* **80**, 489-502.
- Whitelegge, J.P.** (2003). Thylakoid membrane proteomics. *Photosynth Res* **78**, 265-277.
- Wienk, H.L., Wechselberger, R.W., Czisch, M., and de Kruijff, B.** (2000). Structure, dynamics, and insertion of a chloroplast targeting peptide in mixed micelles. *Biochemistry* **39**, 8219-8227.
- Wu, C., Seibert, F.S., and Ko, K.** (1994). Identification of chloroplast envelope proteins in close physical proximity to a partially translocated chimeric precursor protein. *J. Biol. Chem.* **269**, 32264-32271.
- Young, M.E., Keegstra, K., and Froehlich, J.E.** (1999). GTP promotes the formation of early-import intermediates but is not required during the translocation step of protein import into chloroplasts. *Plant Physiol.* **121**, 237-244.
- Zhang, X.P., and Glaser, E.** (2002). Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone. *Trends Plant Sci.* **7**, 14-21.
- Zhong, R., Wan, J., Jin, R., and Lamppa, G.** (2003). A pea antisense gene for the chloroplast stromal processing peptidase yields seedling lethals in *Arabidopsis*: survivors show defective GFP import *in vivo*. *Plant J.* **34**, 802-812.
- Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L., and Gruissem, W.** (2004). GENEVESTIGATOR. *Arabidopsis* microarray database and analysis toolbox. *Plant Physiol.* **136**, 2621-2632.