9-2009

Determinants for Stop-Transfer and Post-Import Pathways for Protein Targeting to the Chloroplast Inner Envelope Membrane

Antonio Americo Barbosa Viana
University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/open_access_dissertations

Part of the Biology Commons

Recommended Citation
Viana, Antonio Americo Barbosa, "Determinants for Stop-Transfer and Post-Import Pathways for Protein Targeting to the Chloroplast Inner Envelope Membrane" (2009). Open Access Dissertations. 137. https://doi.org/10.7275/1079218 https://scholarworks.umass.edu/open_access_dissertations/137

This Open Access Dissertation is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
DETERMINANTS FOR STOP-TRANSFER AND POST-IMPORT PATHWAYS
FOR PROTEIN TARGETING TO THE CHLOROPLAST INNER ENVELOPE
MEMBRANE

A Dissertation Presented

by

Antonio Americo Barbosa Viana

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of
DOCTOR OF PHILOSOPHY

SEPTEMBER 2009

Molecular and Cellular Biology
DETERMINANTS FOR STOP-TRANSFER AND POST-IMPORT PATHWAYS FOR PROTEIN TARGETING TO THE CHLOROPLAST INNER ENVELOPE MEMBRANE

A Dissertation Presented

by

ANTONIO AMERICO BARBOSA VIANA

Approved as to style and content by:

_________________________________________
Danny J. Schnell, chair

_________________________________________
Jennifer Normanly, member

_________________________________________
David Gross, member

_________________________________________
Om Parkash, member

_________________________________________
David Gross, Graduate Program Director
Molecular and Cellular Biology
To my family, for all their love and support.
ACKNOWLEDGEMENTS

I would like to start by thanking my adviser Dr. Danny J. Schnell. During these years, he has consistently shown me not only how to find the way to the answers, but more importantly, to ask the ‘right’ questions and then seek the answers. He has directed me in understanding how to look to a problem and devise the methods and different ways to go about finding the solution. He taught me to elaborate hypotheses, to keep my eyes open to the results and explore the possibilities of critical interpretation. He always made himself available to my questions, and was ready to balance the stress of the laboratory with his patience and wisdom. For several times I rushed into his office unannounced and he was glad to stop his activities to talk about my data. I am very glad to have had Dr. Schnell by my side during this challenging period of my life.

I am also very thankful for Dr. Jennifer Normanly, Dr. David Gross and Dr. Om Parkash, for accepting to be part of my dissertation committee.

I am very thankful for the MCB Ph.D. Program and everyone involved with its success, especially Doreen Fifield and Mary Bell, who were the ones who received me here and showed a great deal of knowledge of all the process and patience to go through it with each one of us (especially the international students). The same goes to the BMB Department, in the name of Angelica LeBoeuf and Carol Ryan. They all did a great work taking care of everything else so I could focus on my research.

Thanks to my friends and colleagues, current and former members of the Schnell lab: Fei Wang, Shu-Zon Wu, Sungwon Han, Alec Derian, Laura Green, Jeonghwa Lee, Hitoshi Inoue and especially Caleb Rounds and Ming Li, for the long discussions, for
sharing frustrations and happiness and for the closer relationships that developed over the
course of the years. All of you contributed to my development as a person and scientist.

I also am profoundly grateful for all my friends; those that I left in Brazil that showed
their support through e-mails and phone calls, and those I made here in Amherst. Most of
them are in the MCB program and it would be unfair to put into a short list. Thanks all of
you ‘MCBers’.

I must also thank my Brazilian friends that had also left their home country and
became a big part of my life here. We are kind of a family, and the support they have
given me can not be put in words.

I have also to express my profound gratitude for my parents and sisters for their
unconditional support. Without their constant support and confidence, I would not have
got this far. Thanks for all the love.

None of this could ever be done without the unconditional love and support of my
wife Paula. She accepted the challenge of coming to a foreign country and dealing with
all the difficulties that would arise to be my very pillar of support. And what a pillar she
has been throughout the years! She has shown extreme devotion and love, being by my
side on the most difficult moments, as well as the most pleasant ones. She has also given
me the most valuable ‘gift’ a man can ever get, and we named her Julia. Whenever the
challenge seemed overwhelming, looking at them gave me the strength I needed to go on.
There are no words that can describe how grateful I am for having both of you in my life.

Finally, but not less important, I would like to express my gratitude to
CAPES/Fulbright (Brazil) and NIH, for funding this research.
ABSTRACT

DETERMINANTS FOR STOP-TRANSFER AND POST-IMPORT PATHWAYS
FOR PROTEIN TARGETING TO THE CHLOROPLAST INNER ENVELOPE
MEMBRANE

SEPTEMBER 2009

ANTONIO AMERICO BARBOSA VIANA, B.S., UNIVERSITY OF BRASILIA,
BRAZIL
M.S., UNIVERSITY OF BRASILIA, BRAZIL
Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Danny J. Schnell

Chloroplast biogenesis relies on the import of thousands of nuclear encoded proteins into the organelle and proper sorting to their sub-organelar compartment. The majority of nucleus-encoded chloroplast proteins are synthesized in the cytoplasm and imported into the organelle via the Toc-Tic translocation systems of the chloroplast envelope. In many cases, these proteins are further targeted to subcompartments of the organelle (e.g. the thylakoid membrane and lumen or inner envelope membrane) by additional targeting systems that function downstream of the import apparatus. The inner envelope membrane (IEM) plays key roles in controlling metabolite transport between the organelle and cytoplasm, and is the major site of lipid and membrane biogenesis within the organelle. In contrast to the protein import and thylakoid targeting systems, our knowledge of the pathways and molecular mechanisms of protein targeting and integration at the IEM are very limited. Previous reports have led to the conclusion that IEM proteins are transferred
to the IEM during protein import via a stop-transfer mechanism. Recent studies have shown that at least two components of the Tic machinery (AtTic40 and AtTic110) are completely imported into the stroma and then re-inserted into the IEM in a post-import mechanism. This led me to investigate the mechanisms and pathways involved in the integration of chloroplast IEM proteins in more detail. I selected candidates (AtTic40 for post-import and IEP37 for stop-transfer) that are predicted to have only one membrane-spanning helix and adopt the same IEM topology to facilitate my analysis. My studies confirm the existence of both stop-transfer and post-import mechanisms of IEM protein targeting. Furthermore, I conclude that the IEP37 transmembrane domain (TMD) is a stop-transfer signal and is able of diverting AtTic40 to this pathway in the absence of AtTic40 IEM targeting information. Moreover, the IEP37 TMD also functions as a topology determinant. I also show that the AtTic40 targeting signals are context dependent, with evidence that in the absence of specific information in the appropriate context, the AtTic40 TMD behaves as a stop-transfer signal. This is an indication that the stop-transfer pathway is the default mechanism of protein insertion in the IEM.
# TABLE OF CONTENTS

| Section                                           | Page |
|---------------------------------------------------|------|
| ACKNOWLEDGEMENTS                                  | v    |
| ABSTRACT                                          | vii  |
| LIST OF TABLES                                    | x    |
| LIST OF FIGURES                                   | xi   |
| CHAPTER                                           |      |
| 1. CHLOROPLAST EVOLUTION AND PROTEIN TARGETING    | 1    |
| 2. INNER ENVELOPE MEMBRANE SORTING IN CHLOROPLASTS| 17   |
| 3. CONCLUSIONS AND PERSPECTIVES                   | 56   |
| 4. EXPERIMENTAL PROCEDURES                        | 64   |
| BIBLIOGRAPHY                                      | 72   |
# LIST OF TABLES

| Table | Description                                                                 | Page |
|-------|-----------------------------------------------------------------------------|------|
| 3.1   | Analysis of TMD properties of known IEM resident proteins                    | 59   |
| 4.1   | List and nucleotide sequence of primers used in deletion, hybrid and substitution mutants, and reaction steps to obtain each construct | 69   |
| Figure | Description |
|--------|-------------|
| 1.1    | Model for the acquisition of the photosynthetic plastid by the Plantaee ancestor |
| 1.2    | Chloroplast structures and envelope functions |
| 1.3    | Major players in the Toc-Tic complex |
| 1.4    | Overview of the two possible pathways for membrane insertion of IEM targeted proteins in chloroplasts |
| 2.1    | Amino acid sequences and membrane topology of IEP37 and AtTic40 |
| 2.2    | IEP37 is a stop-transfer substrate |
| 2.3    | AtTic40 is a post-import substrate |
| 2.4    | Schematic diagram of the constructs used in this study |
| 2.5    | IEP37ΔTM is targeted to isolated chloroplasts and the interaction with the IEM can be disrupted by alkaline extraction |
| 2.6    | The IEP37 TMD targets SSU to the IEM via the stop-transfer pathway |
| 2.7    | The IEP37 TMD targets AtTic40 to the IEM via the stop-transfer pathway |
| 2.8    | Immunoblot of protease treated isolated inside-out chloroplasts IEM vesicles with increasing concentrations of thermolysin |
| 2.9    | IEP37 TMD dictates membrane topology of imported proteins |
| 2.10   | The AtTic40 TMD targets IEP37 to the IEM via the stop-transfer pathway |
| 2.11   | The AtTic40 TMD targets IEP37 to the IEM via the stop-transfer pathway even in the presence of the S/P-rich domain |
| 2.12   | The functionality of the AtTic40 post-import targeting signals is context-dependent |
| 2.13   | The AtTic40 TMD is capable of adopting both orientations in relation to the IEM |
CHAPTER 1

CHLOROPLAST EVOLUTION AND PROTEIN TARGETING

1.1 Chloroplast origin and evolution

A major event in evolution was the development of oxygenic photosynthesis by prokaryotic cyanobacteria around 3 to 3.5 billion years ago. It enriched the planet’s atmosphere with oxygen and generated a massive geochemical transformation, which in turn became the driving force for the striking diversification of O₂-dependent life (Gould, 2008).

In an event referred to as endosymbiosis, formalized by Lynn Margulis (Sagan, 1967), a ‘Plantae’ ancestor cell engulfed and maintained the photosynthetic cyanobacteria, rather than digest it in the food vacuole (Reyes-Prieto et al., 2007). The uptake of this cyanobacteria is known as primary endosymbiosis (figure 1.1) and is considered to be the launch-point of eukaryotic photosynthesis because all canonical plastids appear to be derived, directly or indirectly, from this event (Archibald, 2009).

This relationship, as any kind of symbiosis, came with benefits for both cell types. Whereas the cyanobacterium provided the host cell with valuable resources in the form of carbohydrates from photosynthesis, the host cell provided environmental protection to the newly acquired plastid. Moreover, the relationship became even more interdependent as most of the plastid genome was transferred to the host cell nucleus. This event was a circumvention of some problems generated with the relationship, such as higher tendency
Figure 1.1. Model for the acquisition of the photosynthetic plastid by the Plantae ancestor. In the majority of the occurrences, the cyanobacteria would be digested, but in very rare cases, it might have been retained. Modified from (Reyes-Prieto et al., 2007).
for oxidative damage to the plastid DNA and absence of sexual recombination, which would greatly increase the accumulation of deleterious mutations (Lynch and Blanchard, 1998; Martin and Herrmann, 1998).

As a consequence of endosymbiosis, some genes once encoded and expressed by the bacterium were lost, and others were transferred to the host cell’s nucleus and incorporated into its genome (Timmis et al., 2004). Hence, this close relationship raised some issues that had to be solved in order to be beneficial either for the newly integrated organelle and the host cell. One major issue is that the cellular ribosomal machinery in the cytosol now translates the organelle’s proteins that are encoded by the nuclear genome. Therefore a mechanism needed to evolve in order to target these proteins back into the organelle, ensuring that they are precisely targeted and able to perform their specific functions.

In order for this to happen, these cells developed an intricate targeting system that relies on signals present in these proteins for their proper organelle localization (Blobel et al., 1979). Along with the evolution of these targeting signals, mechanisms for decoding the signals evolved in conjunction in order to guarantee precise protein sorting to and within the organelle.

1.2 Chloroplast sub-compartments and their biogenesis

The chloroplast performs diverse metabolic functions and is a structurally complex organelle, composed of six distinguishable structural and functional units. Three of them are membranes and therefore highly hydrophobic: (i) the outer-envelope membrane
(OEM), (ii) the inner-envelope membrane (IEM), (iii) the thylakoid membrane; and three are aqueous and hydrophilic compartments: (iv) the inter-membrane space (IMS), located between the two envelope membranes, (v) the stroma and (vi) the thylakoid lumen (Jarvis and Robinson, 2004) (figure 1.2).

Around 90% of the proteins that are localized in chloroplasts are encoded by the nuclear genome (Leister, 2003), and thus the accurate targeting of these vast amounts of proteins to the chloroplasts and their further precise sub-organellar localization are of vital importance for cell functionality. Given the high structural complexity of the chloroplast, the targeting of these proteins relies on a very intricate import machinery and sorting system in order to not only reach the chloroplast, but also translocate across its envelope membranes and be properly sorted within the organelle. Protein import and sub-organellar targeting are each mediated by specific targeting signals within the targeted protein. For example, proteins that reside within the thylakoid or IEM contain independent signals for targeting to the organelle and subsequently sorting them to the correct membrane.

1.2.1. Import of nuclear-encoded proteins into the organelle

Chloroplast resident proteins whose genes are nuclear encoded are initially synthesized in the cytosol by the cytosolic translational machinery as precursors (preproteins) and must be imported into the organelle prior to or in conjunction with targeting to the proper sub-organellar compartment.
Figure 1.2. Chloroplast structures and envelope functions. (A) Schematic representation of the structure of a chloroplast. The arrows point to each individual compartment and membrane, whereas the circles indicate the numerous functions performed by the envelope membranes. Modified from (Joyard et al., 1998). (B) Electron micrograph of a chloroplast segment, adapted from (Alberts, 2002), showing its membrane systems. The term ‘grana’ refers to the stacks of thylakoids.
The Toc-Tic (Translocon of the Outer/Inner envelope membrane of Chloroplasts) dependent pathway is the most dominant pathway, and is used by the vast majority of chloroplast targeted proteins (Schnell et al., 1997; Jackson-Constan and Keegstra, 2001; Kessler and Schnell, 2006). The Toc-Tic independent pathways are also referred to as ‘non-canonical’ pathways (Aronsson and Jarvis, 2009). They appear to have evolved to sort a few specialized proteins and, therefore, will not be discussed in detail.

The Toc-Tic pathway requires a machinery that contains more than 20 protein components in Arabidopsis thaliana. Figure 1.3 depicts the major players in the Toc-Tic translocon machinery. All known proteins that use this pathway contain N-terminal cleavable transit peptides (Smeekens et al., 1986; Bruce, 2000). Transit peptides are variable both in length (20-150 amino acid residues) and in primary structure. It has been shown that even though the length and sequence conservation of transit peptides is low, some structural features such as the presence of hydroxylated residues and a lack of acidic residues are notable. Furthermore, transit peptides tend to form random coils in aqueous solutions (von Heijne and Nishikawa, 1991; Bruce, 2000).

After synthesis by the cytosolic translational apparatus, proteins harboring transit peptides are proposed to be kept in the unfolded state by cytosolic chaperones (Hsp70s) (Waegemann, 1990) and 14-3-3 proteins (May and Soll, 2000), which form what is suggested to be a ‘guidance complex’, that directs the precursor proteins to the chloroplast surface.
Figure 1.3. Major players in the Toc-Tic complex. OEM – outer envelope membrane; IMS – inter membrane space; IEM – inner envelope membrane. Modified from (Kessler and Schnell, 2006).
The preprotein is recognized at the chloroplast surface by import receptors that are constituents of the Toc machinery, Toc159 and Toc34. These GTPases act in concert as the primary receptors, and they bind to transit peptides (Kouranov and Schnell, 1997; May and Soll, 1998; Smith et al., 2004). The receptors control access of the preprotein to the Toc machinery via their intrinsic GTPase activities (Kessler and Schnell, 2004, 2006; Wang et al., 2008). Upon initial recognition, the precursor is then delivered to the membrane channel β-barrel protein, Toc75, to initiate translocation (Hinnah et al., 1997; Bolter et al., 1998; Kessler and Schnell, 2006).

After the preprotein crosses the OEM in the unfolded state (Hinnah et al., 2002), it associates with the Tic complex, a translocation machinery that forms a physical contact with the Toc components. The Tic complex physically associates with the Toc complex at membrane ‘contact sites’ (Schnell and Blobel, 1993), which may be mediated by proteins present in the IMS, such as Toc12 and Tic22 (Ma et al., 1996; Kouranov and Schnell, 1997; Becker et al., 2004). The Toc-Tic supercomplex forms a conduit by which translocating proteins can cross both membranes simultaneously, with the assistance of an IMS-resident Hsp70 chaperone (Schnell et al., 1994; Jarvis and Robinson, 2004). It is suggested that the J-domain of Toc12 extends towards the IMS, recruiting Hsp70 to the import machinery and stimulating its ATPase activity which, in turn, aids protein translocation (Becker et al., 2004).

The molecular details of the Tic machinery are less well defined. However, recent studies suggest that preproteins initially engage a 1-MDa inner membrane complex that contains two potential channel components that are proposed to form the Tic core complex, Tic20 and Tic21 (Kikuchi et al., 2009). The preprotein subsequently engages an
abundant inner membrane protein, Tic110. Tic110 interacts directly with preproteins (Inaba et al., 2003), and its abundance relative to other Tic components also suggests its fundamental role in protein translocation (Kouranov et al., 1998). Moreover, it has been shown to be essential for plastid biogenesis, and expression of dominant negative mutants of Tic110 disrupts the assembly of Tic complexes (Inaba et al., 2005).

Tic110 has a large hydrophilic domain that was demonstrated to extend into the stroma and perform multiple functions in protein translocation, including binding with the translocating preprotein as it emerges from the translocon, and serving as a scaffold to recruit stromal chaperones to work as translocation motors and assist in the protein folding. Thus, Tic110 plays a key role in coordinating the stromal events required for protein translocation (Inaba et al., 2003; Inaba et al., 2005). These stromal events also include the recruitment of Tic40 (Stahl et al., 1999; Bedard et al., 2007), a co-chaperone that is also a component of the translocon and is suggested to function in recruiting the chaperone Hsp93 (ClpC), to assist the folding of translocated proteins (Chou et al., 2003; Bedard et al., 2007). Tic40 is proposed to organize the last steps of translocation by regulating the interactions of the translocating preprotein with Tic110 and Hsp93 (Chou et al., 2003). The Tic110, Tic40, Hsp93 complex is proposed to constitute the molecular motor to drive translocation across the membrane.

Finally, upon arrival in the stroma, the transit peptide is cleaved by the stromal processing peptidase (SPP) (Richter and Lamppa, 1999). This cleavage can take place even when the C-terminus of the protein is still bound to the Toc machinery, indicating that the processing occurs as soon as the transit peptide emerges in the stroma (Schnell and Blobel, 1993).
1.2.2. Targeting to the thylakoid membrane

In the case of thylakoid targeted proteins, the targeting signals are bipartite. This means that there is a portion of the N-terminus of the protein that targets it from the cytosol into the organelle, and then another portion that signals the targeting from the stroma to the thylakoid. These processes occur in two independent steps (Smeekens et al., 1986; Ko and Cashmore, 1989). Once the protein has gained access to the stroma, the targeting of proteins to the thylakoid resembles protein export in bacteria because they utilize similar mechanisms (Dalbey and Robinson, 1999; Dalbey and Kuhn, 2000). For that reason, these pathways are referred to as ‘conservative sorting’. Because of this conservation, much is already known about thylakoid targeting from studies in bacterial systems and counterparts in the mitochondria.

Proteins that are destined for the thylakoid membrane or lumen are first targeted to the chloroplast stroma via the Toc-Tic machinery, and additionally employ one of four distinct targeting pathways (Cline and Henry, 1996; Jarvis and Robinson, 2004; Di Cola et al., 2005; Jarvis, 2008). Thylakoid luminal proteins can engage two possible mechanisms: the Sec (secretory) and Tat (twin-arginine) pathways. These substrates contain a bipartite transit sequence that promotes protein sorting via a two-step mechanism. The first N-terminal half of the transit sequence functions as a transit peptide that directs translocation across the envelope via the Toc-Tic complexes into the stroma. There is also a thylakoid luminal targeting domain in the C-terminus of the targeting sequence to direct the proteins into the thylakoid (Schnell, 1998). Proteins that are translocated through the Sec pathway cross the thylakoid membrane through the SecYE
channel in an unfolded state and translocation is energy dependent, with SecA working as the translocation motor, much like in bacterial systems (Cline and Dabney-Smith, 2008).

The Tat pathway substrates, on the other hand, accommodate folded proteins and rely on the thylakoid proton gradient as the energy source. Proteins that are targeted via the membrane-embedded Tat translocase harbor a specialized N-terminal twin-arginine signal peptide containing an SRRXFLK amino acid motif (Robinson and Bolhuis, 2004; Sargent, 2007; Jarvis, 2008).

Thylakoid membrane proteins are targeted by two different pathways. The ‘spontaneous pathway’ inserts proteins that rely solely on a trans-thylakoid ΔpH as an energy source without any soluble factor requirement. The SRP (signal recognition particle) pathway involves GTP-dependent thylakoid membrane insertion, primarily of the highly abundant light harvesting complex proteins. It contains the conserved SRP54 GTPase, but unlike other SRPs, the chloroplast SRP also contains another subunit (SRP43) and lacks RNA (Cline and Dabney-Smith, 2008). This pathway is aided by Alb3, a chloroplast homolog of the mitochondrial Oxa1 and bacterial YidC insertase. The proteins targeted via both pathways do not contain cleavable targeting signals (Jarvis, 2008), and the information to mediate membrane insertion is often in the transmembrane domains (TMD) of such proteins.
1.2.3. Targeting to the envelope membranes

The chloroplast envelope is composed of two membranes, the OEM and IEM. It is the physical barrier that separates the organelle from the cytosol and the only membrane structure that is conserved in different types of plastids. The lipid composition of the envelope membranes is unique, containing primarily galactolipids and sulfolipids with additional lipid derivatives (glycerolipids, terpenoids, carotenoids and chlorophyll precursors) that perform numerous biochemical functions and also are converted into diverse signaling compounds.

The OEM harbors the Toc complex, playing a major role in protein translocation. It had been proposed to be freely permeable to small molecules due to the presence of porins. However, with the discovery of several β-sheet proteins with channel-like activities such as OEP16, OEP24 and OEP21, it is suggested that it participates in the regulated transport of metabolites including amino acids, inorganic phosphate, triose phosphate and hexose phosphates (Weber et al., 2005). Most of the proteins targeted to the outer envelope do not contain a cleavable transit peptide (with the exception of Toc75) and there is evidence for both spontaneous and Toc-dependent membrane insertion (Jarvis, 2008).

The IEM is the selectively permeable barrier between the cytosol and the organelle stroma. It contains proteins involved in chlorophyll and plastoquinone biosynthesis (specifically in the last steps of α-tocopherol and plastoquinone-9 biosynthesis) (Soll et al., 1985), and translocators of products of chloroplast metabolism (e.g. the triose
phosphate/3-phosphoglycerate phosphate translocator (TPT)), which are of fundamental importance for photosynthesis and therefore cell physiology.

The IEM also harbors the type A monogalactosyldiacylglycerol (MGDG) synthase, fundamental in the biogenesis of this lipid, which is the main glycerolipid of this membrane and essential for the biogenesis of the thylakoid membrane (Kobayashi et al., 2009). As such it serves as the origin of both the thylakoid and OEM constituents. The IEM also participates in the coordination of plastid and nuclear gene expression by interacting with the plastid transcription and translation apparatus and by importing nuclear-encoded proteins. Furthermore, the IEM also accommodates the Tic machinery, a key player in protein translocation across the chloroplast envelopes. Therefore, the major roles of the IEM in the organelle and ultimately in plant physiology are unquestionable.

Despite the importance of the envelope in cellular metabolism, signaling and metabolite transport, much less is known about its biogenesis relative to thylakoids, probably due to the lack of a bacterial membrane counterpart. Recent evidence in the literature points to the conclusion that IEM targeted proteins can achieve their destination by at least two possible pathways: the stop-transfer and the post-import pathway (Lubeck et al., 1997; Li and Schnell, 2006; Firlej-Kwoka et al., 2008) (figure 1.4).

The post-import pathway has been recently described (Lubeck et al., 1997; Li and Schnell, 2006), and proteins that are targeted to the IEM through this pathway are imported via the Toc-Tic machinery (figure 1.4).
Figure 1.4. Overview of the two possible pathways for membrane insertion of IEM targeted proteins in chloroplasts.
After the completion of import, the transit peptide is processed by the SPP and the protein can be detected in the soluble phase of the stroma as a soluble intermediate. Subsequently, the protein is inserted in the membrane by a yet unknown machinery, that according to recent evidence, might involve Tic40 (Chiu and Li, 2008). The main feature of this pathway is that the TMD of the protein is translocated across the envelope first, the protein exists as a soluble intermediate in the stroma, and IEM insertion is achieved in an independent downstream event.

The stop-transfer pathway, on the other hand, is widely assumed to be the mechanism of insertion for proteins that do not have a soluble intermediate during its import (figure 1.4). It is known to be an essential mechanism for protein insertion in the membrane during ER co-translational translocation and in bacterial and mitochondrial inner membrane protein insertion (Blobel et al., 1979; Rapoport, 1986; van Loon, 1987; Meier et al., 2005). This process implicates a stop-transfer region that is usually the TMD itself, which halts the threading of the unfolded protein through the translocation channel and is transferred laterally to the lipid bilayer.

1.3 Goal of my research project

There are several fundamental aspects of IEM biogenesis that are still not understood. The intrinsic protein signals that define IEM targeting are yet to be defined. Some TMDs function as stop-transfer signal whereas other TMDs are translocated across the envelope, either to be targeted back to the IEM via the post-import pathway or to be further targeted to the thylakoid membrane. The features that differentiate these TMDs and the
mechanism by which the translocons distinguish between them are still unknown. Moreover, in a more general point of view, the reasons why these different pathways for IEM protein targeting have evolved and what is the importance of each of them in chloroplast biogenesis are questions that remain to be answered. In an effort to begin to understand IEM protein targeting and insertion in more detail, I focused my research in examining the targeting determinants for these two pathways.

1.4 Experimental strategy

There are several aspects that complicate the studies on the biogenesis of the IEM, what may account for the little current information about its dynamics. A major issue is that the vast majority of the known IEM proteins contain mult spanning helices, which makes it difficult to manipulate each TM helix separately and address questions of function and topology. Therefore, I aimed to find proteins that represent the two pathways (stop-transfer and post-import) that are comparable to each other, i.e., contain only one transmembrane helix and attain the same membrane topology.

With these substrates in hand, I used a deletion and substitution approach in which I swapped TMDs and known targeting domains to address the functions of these sequences in the pathway determination, membrane integration and, ultimately, membrane topology.
CHAPTER 2
INNER ENVELOPE MEMBRANE SORTING IN CHLOROPLASTS

2.1 Introduction

The IEM plays central roles in the physiology of the chloroplast and the plant cell. In addition to mediating plastid-cytosolic transport, the IEM is the site of key steps in fatty acid and membrane lipid synthesis and is the origin of all internal plastid membrane structures, including thylakoids. This places the IEM at the fulcrum of membrane biogenesis in the organelle. Despite its central importance, many aspects of IEM biogenesis remain undefined. In particular, the mechanism by which nuclear-encoded IEM proteins are targeted and inserted into the membrane has only recently received attention. In the recent years, evidence in the literature has pointed to at least two possible pathways for IEM translocation/insertion. Lubeck and colleagues (Lubeck et al., 1997) have shown that the major component of the IEM translocation apparatus AtTic110 achieves its IEM localization through a soluble intermediate, i.e., the protein is primarily targeted to the stroma, where it exists for a certain amount of time as a soluble intermediate, and by a mechanism still under investigation, is re-inserted in the IEM and attains its proper topology and folding. A few years later, Li and colleagues (Li and Schnell, 2006) demonstrated that another component of the import apparatus, AtTic40, is also targeted to the IEM by the same pathway. This pathway is known as the ‘post-import’ pathway. The functionality of this pathway was further confirmed by in vivo studies in which the AtTic40 gene was introduced into the chloroplast genome. Pre-
AtTic40 was targeted and successfully inserted into the IEM when expressed from the plastid genome and synthesized on plastid ribosomes (Singh et al., 2008). This was a clear demonstration that the chloroplast indeed possesses a mechanism to target proteins that are produced in the stroma to the IEM, and this mechanism is entirely independent of the envelope translocation process. In other words, membrane insertion can be achieved after protein translocation/import across the chloroplast envelope.

To date, the knowledge about the players in this pathway is very limited. Nevertheless, it has been shown that proteinaceous components of the IEM are required for reinsertion of AtTic40, and exogenous ATP or GTP are not necessary (Li and Schnell, 2006). Moreover, Vojta and colleagues (Vojta et al., 2007) have shown that ATP is necessary for reinsertion of AtTic110 and suggested the participation of the stromal chaperone Hsp93 for its supposed interaction with AtTic110 as indicated by immunoprecipitation studies. More recently, the involvement of AtTic40 on its own and in AtTic110 reinsertion was demonstrated by importing these substrates into AtTic40-null mutant chloroplasts (Chiu and Li, 2008). Even though the detailed mechanism of this pathway remains to be defined, it is suggested that AtTic40 and AtTic110 follow the same pathway.

Less is known about the dynamics of import of other IEM proteins. It is widely assumed that most of the proteins that are imported without the appearance of a soluble intermediate are targeted via the stop-transfer pathway, in a similar fashion to what occurs for membrane protein translocation in the ER and mitochondria. In a screen of the import profiles of several IEM proteins (Firlej-Kwoka et al., 2008), none of the analyzed substrates presented a soluble intermediate form. This is in accordance to what was
shown in a previous study of the Arc6 protein (Vitha et al., 2003). These substrates probably diffuse laterally to the inner membrane concomitantly with translocation, in a mechanism by which a stop-transfer sequence contained in the protein interrupts translocation and orients the polypeptide in the lipid bilayer, as described for the membrane insertion of endoplasmic reticulum and plasma membrane proteins (Blobel et al., 1979), presumably by a lateral opening of the translocation channel.

A few studies were also made in the import dynamics of polytopic IEM proteins, and the most studied example is the triose phosphate translocator (TPT), a 7-helix IEM protein. The TPT is imported into chloroplasts via the Toc-Tic supercomplex and inserted in the IEM (Flugge et al., 1989). A few years later, it was demonstrated that IEM targeting information in the TPT resides in the mature form of the protein, because the transit peptide alone functions as a stromal targeting signal. Moreover, the fusion of the TPT N-terminus containing its transit peptide and first TMD was sufficient to target the small subunit of Rubisco (SSU) to the IEM, (Knight and Gray, 1995) presumably via the stop-transfer pathway. This information led to the hypothesis that the signal that may direct the stop-transfer pathway resides in the TMDs, rather than elsewhere in the mature protein.

Therefore, the goal of this study is to examine the molecular basis for stop-transfer vs. post-import pathways by exploring the signals that determine which pathway is utilized. As a first step, I wanted to identify proteins with similar characteristics that utilize the two pathways. To simplify the analyses, I selected proteins with (i) a single TMD, and (ii) the same topology in relation to the IEM (C-terminus in the stroma, N-terminus in the IMS). AtTic40 is a well-characterized post-import substrate that was shown previously to
fulfill these criteria (Li and Schnell, 2006; Tripp et al., 2007; Chiu and Li, 2008). The AtTic40 protein (figure 2.1A) was studied in some detail in the past few decades, including a few studies on its function, membrane topology and insertion dynamics. It was demonstrated to be localized to the IEM (Stahl et al., 1999) and further studies have shown that AtTic40 has a large C-terminal hydrophilic domain that extends in the stroma (figure 2.1B; Chou et al., 2003). Null mutants of AtTic40 are very pale green, although not lethal, and are defective in chloroplast biogenesis yielding chloroplasts with reduced import rates. AtTic40 was also shown to be part of the translocon machinery. Due to the presence of a tetratricopeptide repeat (TPR) domain followed by a domain with similarities with Sti1p/Hop and Hip co-chaperone proteins (figure 2.1A), it is referred to as a chloroplast co-chaperone that aids protein translocation across the IEM (Chou et al., 2003).

To identify a stop-transfer substrate protein that would be comparable to AtTic40, I performed a systematic search in the ‘Plant Membrane Protein Database – ARAMEMNON’ (www.aramemnon.botanik.uni-koeln.de) (Schwacke R, 2003) and in the literature for inner envelope proteins that contain one single pass TMD. The vast majority of IEM proteins are predicted to be multispans, with the few exceptions of Arc6 (Vitha et al., 2003), H17 (Ferro et al., 2003; Firlej-Kwoka et al., 2008), IEP37 (Motohashi et al., 2003) and AtTic40 (Li and Schnell, 2006; Tripp et al., 2007).

Based on the proteomic studies of several IEM proteins (Ferro et al., 2003), HP17 has a predicted topology (N-terminus in the stroma) that does not match with the criteria I established and it was disregarded as a potential candidate. The same also applies to Arc6, as demonstrated by trypsin accessibility (Vitha et al., 2003). IEP37 (figure 2.1A) was
then selected as a potential candidate because the predictions of its structure and topology fit both criteria (figure 2.1B). Therefore, in this study, I used IEP37 and AtTic40 as model substrates for chloroplast import and insertion dynamics.

The IEP37 protein was first identified as a 37-kDa chloroplast envelope protein (Dreses-Werringloer et al., 1991), and later became also known as APG1 – ‘Albino or Pale Green mutant 1’, when it was found in an albino and pale green A. thaliana mutant screen (Motohashi et al., 2003). Moreover, the same study showed that this mutant’s chloroplasts have decreased levels of chlorophyll and lack plastoquinone. IEP37 also shows partial similarity with the S-adenosyl-dependent methyltransferase, and therefore the APG1 protein (IEP37) is suggested to be involved in plastoquinone biosynthesis and proposed to contain only one C-terminal TMD (figure 2.1). It is the second most abundant protein in chloroplast envelopes, estimated to account for about 5-10% of envelope total protein (Block et al., 1991). In a recent study, IEP37 was shown to have its hydrophilic bulky N-terminus extending to the IMS, being anchored by its C-terminal TM helix to the IEM (Singh et al., 2008) (figure 2.1B).
Figure 2.1. (A) Amino acid sequences of IEP37 and AtTic40. The arrowheads point to sites of transit peptide cleavage. TMD – transmembrane domains. The roman numerals I, II and III represent consensus sequences for S-adenosyl-methionine-dependent methyltransferases (Motohashi et al., 2003). The IEP37 SAM like domain was defined by a CDD search (Marchler-Bauer et al., 2009) in the NCBI Blast website. The S/P-rich domain and TMD in AtTic40 are described in (Tripp et al., 2007) and the TPR and Sti/Hop/Hip domains were identified in an alignment made by (Chou et al., 2003). (B) Schematic representation of the membrane topologies of IEP37 and AtTic40 (Singh et al., 2008).
2.2 Results

2.2.1 IEP37 utilizes the stop-transfer pathway for IEM insertion

In order to better characterize the stop-transfer pathway, I first determined the import profile of pre-IEP37 (figure 2.2A). In vitro translated \(^{35}\)S-Met labeled pre-IEP37 was incubated with isolated chloroplasts for 30 minutes at 20°C. Even though the default chloroplast import reaction temperature is 26°C (Smith et al., 2003), this reduced import temperature (20°C) has been shown in our lab to be optimal to accumulate the AtTic40 soluble intermediate. After import, equivalent samples from the reaction were separated and treated in the presence or absence of thermolysin to remove unimported and surface exposed material. The samples were subsequently fractionated into membrane (P) and soluble (S) fractions using osmotic lysis and alkaline carbonate extraction (0.2 M Na\(_2\)CO\(_3\), pH 11.5), as indicated in figure 2.2B.

As shown in figure 2.2B, pre-IEP37 is imported and processed to its mature form (compare lanes 1 and 2). Mature IEP37 is insensitive to thermolysin treatment, demonstrating that it is fully imported into the organelle (compare lanes 2 and 3). IEP37 fractionates exclusively with the membrane fraction in the presence or absence of alkaline carbonate (pH 11.5) demonstrating that it is indeed integrated into the IEM (figure 2.2B, compare lanes 4 and 6 with 5 and 7, and graph). A time course was performed in an attempt to observe a possible soluble intermediate (figure 2.2C). Even after only 2 minutes of import, no significant amount of soluble form can be detected, suggesting that this protein inserts into the IEM without a soluble intermediate.
Figure 2.2. IEP37 is a stop-transfer substrate. (A) Schematic of the pre-IEP37 protein. (B) The [S\(^{35}\)]-Met construct in (A) was imported into isolated chloroplasts for 30 minutes and treated as indicated. The graph represents the quantification of lanes 4, 5, 6 and 7. ‘Control’ and ‘pH11.5’ on graph indicates osmotic lysis and alkaline extraction, respectively. (C) The construct was imported at 20 °C for 2, 5 and 30 minutes, treated with thermolysin and separated into membrane (P) and supernatant (S) fractions by alkaline extraction. (D) The radiolabelled construct was imported for 5 minutes at 20°C, the reaction stopped, treated with thermolysin and resumed in the presence of 5mM ATP (chase). At the indicated time points, equivalent fractions were collected and separated into membrane (P) and supernatant (S) fractions by osmotic lysis. (E) Alkaline extraction of the last chase time point. Thermolysin (200 µg/ml) treatments were performed on ice for 30 minutes.
To eliminate the possibility that IEP37 targeting would involve a post-import intermediate that simply remained associated with the IEM, but was not membrane integrated, I performed a chase assay, in which the import reaction proceeded for 5 minutes at 20 °C to accumulate a potential intermediate. The import reaction was stopped, the chloroplasts were treated with thermolysin to remove any unimported protein, and import was resumed (chase) in the presence of 5 mM ATP at 26 °C. Samples were taken at the beginning of the chase reaction (time = 0), at 5 and at 60 minutes, and subsequently fractionated into membrane (P) and soluble (S) fractions. The last time point (after the 60-minute chase) was also alkaline extracted (pH 11.5) to certify the proper integration of the substrate in the membrane. An identical assay was shown to accumulate AtTic40’s soluble intermediate (Li and Schnell, 2006). The chase reaction confirms the absence of a soluble form during the import of IEP37 (figure 2.2D lanes 7, 10 and 13). The graph in figure 2.2D shows that the levels of membrane-bound mature IEP37 remain constant with time, indicating that all detectable protein was already membrane integrated at the beginning of the chase. The quantification of the soluble form indicates that it represents less than 5% of total IEP37 at all times. Even after alkaline extraction, the proportion of membrane integrated IEP37 remains nearly the same as at the beginning of the chase (compare fig 2.2D and E). These data indicate that IEP37 import does not involve a soluble form, which is in agreement with the data presented by (Firlej-Kwoka et al., 2008). Therefore, I conclude that membrane insertion of IEP37 occurs during translocation via a stop-transfer translocation/insertion.

For comparison, I performed similar import experiments with pre-AtTic40 (figure 2.3A). A standard 30-minute import reveals that both mature AtTic40 and an
intermediate sized form (int-AtTic40) are present (figure 2.3B, lane 2). Int-AtTic40 is generated by processing of the transit peptide by the stromal processing peptidase. Mature AtTic40, on the other hand, is generated by a processing event by an unknown protease at the inner membrane during post-import targeting (Li and Schnell, 2006). The intermediate and mature forms are thermolysin insensitive (compare figure 2.3B lanes 2 and 3). As expected, approximately 45% of the intermediate-sized protein is found in the soluble form after alkaline extraction (figure 2.3B, lanes 6, 7; graph on the right). In a chase experiment, the soluble intermediate-sized form can be detected in the beginning of the chase (time = 0 min) (figure 2.3C, lane 7) and its intensity decreases progressively as the chase proceeds (figure 2.3C, compare lanes 7, 10 and 13). At the same time, the mature membrane associated form band increases to the same proportion (figure 2.3C, compare lanes 6, 9 and 12), indicating direct conversion from the soluble intermediate to the mature membrane integrated form. The graph in panel C represents the conversion of the total detected intermediate to mature form, as another confirmation that the conversion takes place. In figure 2.3D, alkaline extraction of the 60-minute chase reveals that most of the protein is indeed membrane integrated (~80%). Taken together these data confirm previous observations (Li and Schnell, 2006) and are consistent with the protein being fully imported into the stroma, processed to the intermediate form and inserted in the IEM and processed to the mature form. In conclusion, AtTic40 is targeted to the IEM via the post-import pathway.
Figure 2.3. AtTic40 is a post-import substrate. (A) Schematic of the pre-AtTic40 protein. (B) The [S\textsuperscript{35}]\textsuperscript{-Met construct in (A) was imported into isolated chloroplasts for 30 minutes and treated as indicated. The graph represents the quantification of lanes 4, 5, 6 and 7 for either the mature or the intermediate-sized forms. 'Control' and 'pH11.5' on graphs indicate osmotic lysis and alkaline extraction, respectively. (C) The radiolabelled construct was imported for 5 minutes at 20 \textdegree C, the reaction stopped, treated with thermolysin and resumed in the presence of 5mM ATP (chase). At the indicated time points, equivalent fraction were collected and separated into membrane (P) and supernatant (S) fractions by osmotic lysis. (D) Alkaline extraction of the last chase time point. Thermolysin (200 \mu g/\mu l) treatments were performed on ice for 30 minutes.
2.2.2 The IEP37 TMD is necessary and sufficient for membrane integration via the stop-transfer pathway

With two model substrates for both the stop-transfer and post-import pathways in hand, I set out to define the intrinsic features of the proteins that confer the specificity for the membrane insertion pathway. In order to do so, I generated a series of deletion and substitution mutants (figure 2.4) that were individually and systematically studied further.

As the TMD is the region of the protein that anchors the protein in the membrane, and has been shown to be responsible for IEM targeting of the TPT (Knight and Gray, 1995), I started by deleting the TMD in pre-IEP37 (pre-IEP37ΔTM) to ask if it is a necessary signal for IEM targeting (figure 2.5A). In vitro translated[^35S]-Met pre-IEP37ΔTM was incubated with isolated chloroplasts under import conditions. After 30 minutes of import reaction, most of the imported protein appears in the membrane fraction when simple osmotic lysis of the chloroplasts (2mM EDTA) is used to separate the membrane and soluble fractions (figure 2.5B, lanes 8 and 9). However, upon treatment with alkaline pH, ~95% of all detectable protein is found at the soluble phase (figure 2.5B, compare lanes 10 and 11; graph). The results show that the IEP37 TMD is necessary for proper membrane integration.
Figure 2.4. Schematic diagram of the constructs used in this study. The empty arrowheads represent transit peptide cleavage sites. The numbers indicate the amino acid residue position of the pre-protein.
Figure 2.5. IEP37ΔTM is targeted to isolated chloroplasts and the interaction with the IEM can be disrupted by alkaline extraction. (A) Schematic of the pre-IEP37ΔTM construct. (B) The [S\textsuperscript{35}]-Met construct in (A) was imported into isolated chloroplasts for 30 minutes and treated as indicated. The graph represents the quantification of lanes 10 and 11 in the gel. Thermolysin (200 µg/µl) treatments were performed on ice for 30 minutes.
In order to determine if the IEP37 TMD is sufficient for protein targeting to the IEM, I generated a hybrid construct in which the IEP37 transit peptide and TMD were fused to rubisco small subunit (SSU), a soluble stromal protein. The hybrid construct, named pre-IEP37TP-SSU-IEP37TM (figure 2.6A), was then subjected to a chloroplast in vitro import assay to determine whether it is properly targeted and sorted.

The addition of the IEP37 TP and TMD to SSU was efficient in directing the protein to the organelle (figure 2.6B, lanes 2 and 3) and to the membrane fraction (compare lanes 4, 5, 6 and 7). An import time course revealed no significant soluble intermediate, even at the early time points (fig. 2.6C, lanes 3, 5 and 7). Therefore, the fusion protein also achieved its membrane insertion without the appearance of a soluble intermediate. The mature form is totally resistant to alkaline extraction, revealing that the protein is indeed integrated in the membrane (fig 2.6C lanes 2, 4 and 6; graph).

During the analysis of the pre-IEP37TP-SSU-IEP37TM construct, I observed that the imported, mature protein was degraded with continued incubation under import conditions. This phenomenon was also observed when similar fusions were made to GFP and by placing the IEP37 TMD at the N-terminus of the mature SSU (data not shown). This suggests that non-native proteins might be unstable when targeted to the IEM or exposed to the IMS. Therefore, we chose to further examine the targeting properties of the IEP37 TMD in the context of a native IEM protein.

For this purpose, I replaced the post-import targeting signal of AtTic40 with the IEP37 TMD. Tripp and colleagues (Tripp et al., 2007) showed using AtTic40 deletion mutants that the AtTic40 TMD and an S/P-rich domain adjacent to the TMD are
Figure 2.6. The IEP37 TMD targets SSU to the IEM via the stop-transfer pathway. (A) Schematic of the pre-IEP37TP-SSU-IEP37TM construct. (B) The [S\textsuperscript{35}]-Met construct in (A) was imported into isolated chloroplasts for 30 minutes and treated as indicated. The graph represents the quantification of lanes 4, 5, 6 and 7. ‘Control’ and ‘pH11.5’ on graph indicates osmotic lysis and alkaline extraction, respectively. (C) The radiolabelled construct was imported at 20 °C for 2, 5 and 30 minutes, treated with thermolysin and separated into membrane (P) and supernatant (S) fractions by alkaline extraction. The graph represents the quantification of the gel. Thermolysin (200 µg/µl) treatments were performed on ice for 30 minutes.
necessary and sufficient for membrane integration. Therefore, I deleted this S/P-rich domain along with AtTic40 TMD, and put IEP37 TMD in its place (pre-AtTic40-IEP37TM; figure 2.7A). This construct was subjected to an in vitro import reaction. The results in figure 2.7B show that the construct was efficiently imported as indicated by thermolysin resistance (lane 3) and processed to its intermediate size. Remarkably, the second processing that converts int-AtTic40 to mature AtTic40 did not occur efficiently. This indicates that the N-terminus has reached the stroma, but the second protease could not efficiently reach the second processing site. The deletion of the S/P-rich domain and TMD of AtTic40 is far removed from the second processing site (at Ala76, figure 2.7A) and is unlikely to account for the lack of cleavage. A previous study (Tripp et al., 2007) showed that even with the deletion of residues 77 to 82 or 83 to 88, the second processing occurs efficiently. In the AtTic40-IEP37TM construct, the AtTic40 N-terminus was preserved up to residue 83 (figure 2.7A).

A time course of pre-Tic40-IEP37TM import (figure 2.7C) revealed that the intermediate-sized form (int-Tic40-IEP37TM) accumulated over time in the membrane fractions, with approximately only 10% of the protein detected in the soluble phase after alkaline extraction (figure 2.7C; graph). The extractable component remains essentially constant throughout the course of import, suggesting that it is not an intermediate in the targeting process.

To confirm that this construct is indeed imported through the stop-transfer pathway, I performed a chase experiment (figure 2.7D) and extracted the membrane proteins at the end of the chase (figure 2.7E). The chase reveals no significant soluble free form (figure
Figure 2.7. The IEP37 TMD targets AtTic40 to the IEM via the stop-transfer pathway. (A) Schematic of the pre-AtTic40-IEP37TM construct. (B) The [S\textsuperscript{35}]-Met construct in (A) was imported into isolated chloroplasts for 30 minutes and treated as indicated. The graph represents the quantification of lanes 4, 5, 6 and 7. ‘Control’ and ‘pH11.5’ on graph indicates osmotic lysis and alkaline extraction, respectively. (C) The radiolabelled construct was imported at 20 °C for 2, 5 and 30 minutes, treated with thermolysin and separated into membrane (P) and supernatant (S) fractions by alkaline extraction. The graph represents the quantification of the gel. (D) The radiolabelled construct was imported for 5 minutes at 20 °C, the reaction stopped, treated with thermolysin and resumed in the presence of 5mM ATP (chase). At the indicated time points, equivalent fractions were collected and separated into membrane (P) and supernatant (S) fractions by osmotic lysis. (E) Alkaline extraction of the last chase time point. Thermolysin (200 µg/µl) treatments were performed on ice for 30 minutes.
2.7D lanes 7, 10 and 13; graph), and alkaline extraction of the 60-minute chase time point confirms that the levels of alkaline extractable int-Tic40-IEP37TM remain at ~10% of the total even after the end of the chase (figure 2.7E). This is evidence that the soluble form that appears in low amounts (~10%) under high pH conditions is not converted into a mature form, and is further confirmation that this fusion protein is inserted in the membrane via the stop-transfer pathway.

Therefore, I conclude that the IEP37 TMD is necessary and sufficient for membrane insertion of both a non-native imported protein (SSU) as well as an inner envelope resident protein (AtTic40) via the stop-transfer pathway.

2.2.3 The IEP37 TM helix dictates membrane topology of imported proteins

During protein import into chloroplasts via the Toc-Tic translocon, the N-terminal transit peptide is recognized by receptors at the surface of the chloroplast and translocation proceeds N-terminus first. In the case of a C-terminally anchored protein such as IEP37, with the bulky N-terminus extending into the IMS, the polypeptide must flip within the translocon to attain its correct topology during translocation.

The fact that the AtTic40-IEP37TM construct was imported via the stop-transfer pathway and the second processing was inhibited prompted me to investigate the membrane orientation of this construct. This could account for the lack of accessibility of the protease responsible for the generation of the mature form. IEP37 and AtTic40 each have bulky hydrophilic segments. In the case of AtTic40, a small N-terminal region extends into the IMS with the bulk of the protein, including its C-terminus, in the stroma.
In contrast, the bulk of IEP37 and its N-terminus reside in the IMS (figure 2.1B). The difference in the distribution of the AtTic40 and IEP37 polypeptides across the IEM provided a method to readily assess topology. In isolated, inside-out IEM vesicles treated with protease, AtTic40 is expected to be completely degraded, whereas IEP37 is expected to be largely intact (Singh et al., 2008). To test this assay, I first isolated inside-out IEM vesicles using a standard protocol (Keegstra and Yousif, 1986; Smith et al., 2003), and titrated the amount of protease in order to differentiate the topologies of AtTic40 and IEP37. To probe the resistance/susceptibility of the proteins, I used antibodies against AtTic40 and IEP37. The results shown in figure 2.8 demonstrate that thermolysin is a well suited protease to this assay. In isolated inside-out IEM vesicles, IEP37 is largely protected from protease digestion because the bulk of the protein is located within the vesicle lumen (compare lanes 1-6). With increasing amounts of protease, the IEP37 protease protected band is converted into a slightly smaller polypeptide due to the cleavage of the short C-terminal tail of the protein that is exposed outside of the vesicles. In contrast, AtTic40 is almost completely degraded at the lowest concentration of protease tested (10 µg thermolysin/mg total protein; compare lanes 1 and 2). Both proteins are efficiently degraded if triton X-100 is included in the assay to disrupt membrane integrity (figure 2.8, lanes 7-12). Based on this assay, I determined that the minimum amount of protease to distinguish between AtTic40 and IEP37 distribution is 20 µg thermolysin/mg of total protein (figure 2.8, lanes 3 and 9).

Radiolabelled pre-IEP37 and pre-AtTic40-IEP37TM were imported into isolated purified chloroplasts for 30 minutes and treated with thermolysin to digest anything that was not successfully imported. I then isolated IEM inside-out vesicles using standard
Figure 2.8. Immunoblot of protease treated isolated inside-out chloroplasts IEM vesicles with increasing concentrations of thermolysin, probed with the indicated antibodies. The asterisk indicates a cross-reaction of the anti-AtTic40 antibody that is converted into the double-asterisk indicated band with increasing concentrations of protease. TX-100 indicates the presence of 2% of triton X-100.
procedures from chloroplasts after both import reactions (Keegstra and Yousif, 1986; Smith et al., 2003).

The IEM vesicles isolated from chloroplasts that imported the constructs individually were then treated with 20 µg of thermolysin/mg of total protein in order to reveal the distribution of the imported polypeptides across the membrane. The upper panel in figure 2.9 shows that imported radiolabelled IEP37 is largely resistant to protease digestion (upper panel, lane 3), confirming the immunoblot results for the native protein (figure 2.8). The graph in figure 2.9 reveals that about 80% of the total protease untreated signal is still detectable, and upon membrane solubilization, virtually all IEP37 is digested (upper panel, lane 4).

The replacement of the AtTic40 IEM insertion signal with the IEP37 TMD not only directed this substrate to the membrane through the stop-transfer pathway, but also promoted correct membrane orientation with the N-terminus in the IMS. This was revealed by protease accessibility of the imported protein (lower panel lane 3; graph). In order to acquire this N-out/C-in topology without going through a soluble form and being post-import inserted in the IEM, the protein has to flip during translocation, presumably without leaving the translocon, just as native IEP37. This topology also might account for the inefficient processing of the second cleavage site. In native AtTic40, cleavage appears to occur at the stromal face of the IEM during post-import insertion. Taken together, these results led me to conclude that the IEP37 TMD is a signal not only for the stop-transfer pathway, but also for the determination of its membrane orientation. Moreover, the IEP37 TMD inserts in the IEM in the N-out/C-in orientation only. The schematic drawings in figure 2.9 represent the final topology observed for the two proteins.
Figure 2.9. IEP37 TMD dictates membrane topology of imported proteins. The indicated radiolabelled constructs were imported into isolated chloroplasts for 30 minutes at 26 °C. After import, the chloroplasts were treated with 200 µg/µl thermolysin for 30 minutes on ice and inside-out IEM vesicles were isolated. The vesicles were then treated or not (as indicated) with 20 µg/µl thermolysin in the presence and absence of 2% Triton X-100 (TX-100). On top, IEM inside-out vesicles of chloroplasts pre-imported with pre-IEP37. The graph depicts the densitometric quantification of the gel and the scheme on the right illustrates the proposed mechanism of IEM insertion. On the bottom, pre-AtTic40-IEP37TM was subjected to the same treatment.
2.2.4 The AtTic40 TMD can function as a stop-transfer signal

My next goal was to examine how the IEP37 and AtTic40 insertion signals dictate the stop-transfer and post-import pathways, respectively. The TMD of IEP37 appears to be necessary and sufficient for targeting via the stop-transfer pathway, whereas the stop-transfer pathway utilized by AtTic40 appears to require both a TMD and the presence of the S/P-rich domain upstream of the TMD in AtTic40 (Tripp et al., 2007). This prompted me to investigate whether the TMDs of IEP37 and AtTic40 play active roles in determining the IEM targeting pathway or the pathway is determined primarily by the presence or absence of the S/P-rich domain.

As a first step, I tested whether the AtTic40 TMD could function in IEM targeting in the absence of the S/P-rich domain. To this end, I replaced the IEP37 TMD (residues 307-328) with the AtTic40 TMD (residues 106-131) to generate pre-IEP37-Tic40TM (figure 2.10A). This construct was imported into isolated chloroplasts for 30 minutes and the results are shown in figure 2.10B. After a 30-minute import reaction, most of the imported protein (~90%) co-fractionated with the membrane fraction, and upon alkaline extraction, roughly 75% of the protein remains tightly bound to the membrane. A time course of pre-IEP37-Tic40TM import (figure 2.10C) revealed that the small proportion of IEP37-Tic40TM that is not membrane integrated remains essentially constant as a proportion of total protein from 2 minutes to 30 minutes of import, which is an indicator of the stop-transfer pathway. I performed a chase under the same conditions established
Figure 2.10. The AtTic40 TMD targets IEP37 to the IEM via the stop-transfer pathway. (A) Schematic of the pre-IEP37-Tic40TM construct. (B) The [S\textsuperscript{35}]-Met construct in (A) was imported into isolated chloroplasts for 30 minutes and treated as indicated. The graph represents the quantification of lanes 4, 5, 6 and 7. ‘Control’ and ‘pH11.5’ on graph indicates osmotic lysis and alkaline extraction, respectively. (C) The radiolabelled construct was imported at 20 °C for 2, 5 and 30 minutes, treated with thermolysin and separated into membrane (P) and supernatant (S) fractions by alkaline extraction. The graph represents the quantification of the gel. (D) The radiolabelled construct was imported for 5 minutes at 20 °C, the reaction stopped, treated with thermolysin and resumed in the presence of 5mM ATP (chase). At the indicated time points, equivalent fractions were collected and separated into membrane (P) and supernatant (S) fractions by osmotic lysis. (E) Alkaline extraction of the last chase time point. Thermolysin (200 µg/µl) treatments were performed on ice for 30 minutes. The asterisk indicates a degradation product.
previously, and once again, there were very low levels of soluble IEP37-Tic40TM (~4% of total IEP37-Tic40TM) detected at each time point in the chase (figure 2.10D, lanes 7, 10 and 13; graph). Alkaline extraction of the 60-minute time point after the chase revealed that approximately 30% (Figure 2.10E) of the protein is not integrated. The amount of protein loosely associated with the membrane is slightly higher than that observed at the 5-minute time point in import that was used at the start of the chase (compare the graphs in figures 2.10C and E). This suggests that the small proportion of IEP37-Tic40TM that associates loosely with the IEM does not represent a productive targeting intermediate, but likely corresponds to a small amount of mistargeted protein. Therefore, I conclude that the AtTic40 TMD can function as a stop-transfer signal in the absence of the S/P-rich domain, and is capable of targeting IEP37 to the IEM, albeit with a slightly lower efficiency than the native IEP37 TMD.

To further explore the signals in AtTic40 that may direct the post-import pathway, I generated a fusion in which I replaced the TMD in IEP37 with the AtTic40 S/P-rich domain and TMD. This construct was named pre-IEP37-Tic40SPTM (figure 2.11A). Figure 2.11B shows that pre-IEP37-Tic40SPTM is not only properly targeted to the chloroplasts, but effectively inserted into the membrane, being present in the membrane fraction as shown by alkaline extraction (figure 2.11B, lanes 6 and 7; graph), with less than 20% of total protein in the soluble phase upon alkaline extraction.

A time course of IEP37-Tic40SPTM import (figure 2.11C) also revealed that the proportion of protein found in the soluble phase after alkaline extraction remains essentially constant, at less than 20% of total protein from the beginning of import.
Figure 2.11. The AtTic40 TMD targets IEP37 to the IEM via the stop-transfer pathway even in the presence of the S/P-rich domain. (A) Schematic of the pre-IEP37-Tic40SPTM construct. (B) The [S\textsuperscript{35}]-Met construct in (A) was imported into isolated chloroplasts for 30 minutes and treated as indicated. The graph represents the quantification of lanes 4, 5, 6 and 7. ‘Control’ and ‘pH11.5’ on graph indicates osmotic lysis and alkaline extraction, respectively. (C) The radiolabelled construct was imported at 20 °C for 2, 5 and 30 minutes, treated with thermolysin and separated into membrane (P) and supernatant (S) fractions by alkaline extraction. The graph represents the quantification of the gel. (D) The radiolabelled construct was imported for 5 minutes at 20 °C, the reaction stopped, treated with thermolysin and resumed in the presence of 5mM ATP (chase). At the indicated time points, equivalent fractions were collected and separated into membrane (P) and supernatant (S) fractions by osmotic lysis. (E) Alkaline extraction of the last chase time point. Thermolysin (200 µg/µl) treatments were performed on ice for 30 minutes. The asterisk indicates a degradation product.
The amount of loosely bound protein remains the same at each time point if the chloroplasts from a 5 min import are reisolated and chased (figure 2.11D). Alkaline extraction of the sample from the 60-minute chase (figure 2.11E) reveals that 20% of the total protein is found in the soluble phase. This amount is similar to that observed at the start of the chase (figure 2.11C, 5 min time point), arguing that this form is not a productive targeting intermediate.

In conclusion, I have shown that the AtTic40 TMD in the presence or absence of the S/P-rich domain functions as a stop-transfer signal when inserted into a native IEM protein with similar topology. I therefore conclude that the stop-transfer pathway is the default pathway for chloroplast IEM targeting.

2.2.5 The AtTic40 S/P-rich domain and TMD functionality as a post-import signal is context-dependent

A previous study demonstrated that fusion of the AtTic40 transit peptide, S/P-rich domain and TMD to the N-terminal region of GFP could target the protein to the IEM via a soluble intermediate similar to that observed with native AtTic40 (Tripp et al., 2007). These results are in contrast with those in figure 2.11 in which the S/P-rich domain and TMD of AtTic40 arrested the IEP37-Tic40SPTM fusion on the IEM in the stop-transfer pathway. The major difference between the strategies used in the two studies was the position of the AtTic40 targeting domain within the fusion proteins. The GFP fusion placed the signals at the N-terminus in a context comparable to those found in native AtTic40, whereas the pre-IEP37-Tic40SPTM construct placed the signals at the C-
terminal region of the protein. The position of the AtTic40 signals in the protein might play a role in the function of these signals and ultimately in the pathway determination.

To test this possibility, I generated a construct that contained the AtTic40 S/P-rich domain and TMD at the N-terminus of IEP37. The construct, pre-Tic40N-IEP37ΔTM (figure 2.12A), contained the N-terminus of AtTic40 up to the TMD (including the S/P-rich domain) and was fused with a truncated form of IEP37 lacking its C-terminus TMD. This construct was imported into isolated chloroplasts and the results in figure 2.12B revealed that the protein is properly imported into chloroplasts as shown by thermolysin resistance (lane 3). Interestingly, this construct is processed twice upon import, once to an intermediate size, and again to the mature form. This is similar to native AtTic40 (Li and Schnell, 2006) and figure 2.3). A significant portion of the intermediate (int-Tic40N-IEP37ΔTM) is extractable (>30%) from the membrane at early time points in import (figure 2.12C, compare lanes 2 and 3; graph).

To directly test whether or not int-Tic40N-IEP37ΔTM was a targeting intermediate, I performed a chase experiment (figure 2.12D). The abundance of the int-Tic40N-IEP37ΔTM form decreased during the chase in the same proportion that the mature band increases (figure 2.12D; graph). This observation confirms that int-Tic40N-IEP37ΔTM is directly converted to the mature form. The mature form of Tic40N-IEP37ΔTM is largely insensitive to alkaline extraction from the membrane fraction, demonstrating that it is fully integrated into the membrane in the same proportion as AtTic40 (~80% integration) (compare figures 2.3D and 2.12D). It is interesting to note that int-Tic40N-IEP37ΔTM is peripherally associated with the membrane, and is only released by alkaline extraction.
Figure 2.12. The functionality of the AtTic40 post-import targeting signals is context-dependent. (A) Schematic of the pre-Tic40N-IEP37ΔTM construct. (B) The [S$^{35}$]-Met construct in (A) was imported into isolated chloroplasts for 30 minutes and treated as indicated. The graph represents the quantification of lanes 4, 5, 6 and 7. ‘Control’ and ‘pH11.5’ on graph indicates osmotic lysis and alkaline extraction, respectively. (C) The radiolabelled construct was imported at 20°C for 2, 5 and 30 minutes, treated with thermolysin and separated into membrane (P) and supernatant (S) fractions by alkaline extraction. The graph represents the quantification of the gel. (D) The radiolabelled construct was imported for 5 minutes at 20°C, the reaction stopped, treated with thermolysin and resumed in the presence of 5mM ATP (chase). At the indicated time points, equivalent fractions were collected and separated into membrane (P) and supernatant (S) fractions by osmotic lysis. (E) Alkaline extraction of the last chase time point. Thermolysin (200 µg/µl) treatments were performed on ice for 30 minutes.
This is in contrast to int-AtTic40, which is found largely as a soluble intermediate in the stroma (figure 2.3B and C; (Li and Schnell, 2006)). IEP37ΔTM also remained loosely associated with the membrane even though it lacked a TMD (figure 2.5B lanes 3 and 8), suggesting that IEP37 remains loosely associated with the membrane even when the membrane targeting signals are removed.

In conclusion, my data suggest that the function of the AtTic40 S/P-rich domain and TMD as post-import targeting signals is dependent upon the position of the signal within the polypeptide. In the inappropriate position (e.g. C-terminus), the AtTic40 S/P-rich domain and TMD or the TMD alone are capable of directing IEM insertion, but they function as stop-transfer signals.

### 2.2.6 The AtTic40 TM helix adopts both topologies in relation to the IEM

In addition to promoting membrane insertion, the IEM targeting signals in IEP37 and AtTic40 also dictate the topology of the protein in the membrane. In the case of IEP37, the TMD must act as a stop-transfer signal and trigger the polypeptide to flip during translocation to adopt the proper orientation (N-terminus in the IMS). Therefore, it was of interest to determine the topology of the IEP37 fusions containing the AtTic40 targeting signals. I analyzed the topology of IEP37-Tic40SPTM and AtTic40N-IEP37ΔTM in comparison with AtTic40 using the inside-out IEM vesicles isolation and protease protection assay described previously.

When I replaced IEP37 TMD by AtTic40 TMD (pre-IEP37-Tic40SPTM) (figure 2.13, middle panel), the hybrid IEP37 in the vesicles was protease sensitive, indicating
that the bulk of the protein is exposed at the stromal face of the IEM. This is the reverse topology of native IEP37, indicating that although the AtTic40 segments function as stop-transfer signals in this context, they are unable to cause the polypeptide to flip and assume its native topology. These data also confirm that regions of IEP37 outside of the TMD do not play a direct role in determining topology. The protease sensitivity pattern of IEP37-Tic40SPTM was identical to native AtTic40 (figure 2.13, upper panel), which was used as a control for the C-in/N-out topology.

In the case of AtTic40N-IEP37ΔTM, the polypeptide was completely susceptible to protease digestion and therefore also exposed at the stromal face of the IEM (figure 2.13, lower panel). This confirms that the AtTic40 N-terminus containing its S/P-rich domain and TMD is capable of targeting IEP37 lacking its own TMD to the IEM via the post-import way, and inserting the protein in the membrane in the same topology as AtTic40.
Figure 2.13. The AtTic40 TMD is capable of adopting both orientations in relation to the IEM. The indicated radiolabelled constructs were imported into isolated chloroplasts for 30 minutes at 26 °C. After import, the chloroplasts were treated with 200 µg/µl thermolysin for 30 minutes on ice and inside-out IEM vesicles were isolated. The vesicles were then treated or not (as indicated) with 20 µg/µl thermolysin in the presence and absence of 2% Triton X-100 (TX-100). On top, IEM inside out vesicles of chloroplasts pre-imported with pre-AtTic40. The graph depicts the densitometric quantification of the gel and the scheme on the right illustrates the proposed mechanism of IEM insertion. On the middle and bottom panels, pre-IEP37-Tic40SPTM and pre-Tic40N-IEP37ΔTM, were subjected to the same treatment, respectively.
2.3 Discussion

In order to determine the features that dictate the differentiation of substrates between the post-import and stop-transfer pathway, I first analyzed the role of IEP37 TMD in its own import and membrane insertion. Furthermore, I examined the dynamics of insertion of a stroma targeted protein (SSU) and a native IEM resident protein AtTic40 that is known to adopt the post-import pathway for IEM integration. I confirmed previously reported data that IEP37 import does not involve a soluble intermediate (Firlej-Kwoka et al., 2008), being inserted in a stop-transfer fashion. Moreover, the deletion of its TMD proved that it is necessary for its insertion in the membrane, and the addition of its TMD to SSU was necessary and sufficient to target this hybrid protein to the IEM in the same pathway. Therefore, I conclude that IEP37 TMD is sufficient to drive the native IEP37 protein, an otherwise soluble protein (SSU), and an IEM resident protein, AtTic40, to the IEM through the stop-transfer pathway.

With the knowledge of the necessity and sufficiency of IEP37 TMD for membrane insertion through the stop-transfer pathway, I hypothesized that by providing the appropriate signaling context, it would be possible to divert the insertion pathway by which the proteins integrate into the chloroplast IEM. Therefore, using overlap-extension PCR (Horton et al., 1989; Horton et al., 1993), I generated a series of domain swap constructs containing IEP37 and AtTic40 TMDs (pre-IEP37-Tic40TM, pre-Tic40-IEP37TM) and the AtTic40 S/P-rich domain (pre-IEP37-Tic40SPTM). In the native AtTic40, the S/P-rich domain is present upstream of the TMD (figures 2.1 and 2.4), precisely between the second cleavage site and the TMD, that is, as the name implies, rich in serine and proline residues (10 serines and 5 prolines). However, it has been
reported that only the C-terminal part of this domain is relevant to the IEM insertion of AtTic40 (6 serines and 5 prolines) (Tripp et al., 2007). Therefore, this segment of the S/P-rich domain was included in my analyses (figure 2.4).

The domain swap constructs were subjected to import, chase and topology determination assays, which revealed that the addition of the AtTic40 TMD to IEP37 constructs, regardless of the presence of the S/P-rich domain, does not divert IEP37 to the post-import pathway. Instead, these constructs presented the very same import profile as IEP37 in isolated chloroplasts import followed by a chase, without a detectable and chaseable soluble form. The hybrid IEP37 constructs are successfully imported, processed to the mature form and efficiently inserted into the membrane, as confirmed by alkaline extraction, revealing that the introduction of the signals that were previously shown to be essential for the membrane integration of AtTic40 (Tripp et al., 2007), although necessary for membrane integration, are not sufficient for establishing the post-import pathway when in the context of IEP37 C-terminus.

The AtTic40 construct containing the IEP37 TMD, on the other hand, was diverted from its original post-import to the stop-transfer pathway. This alone is a very revealing observation, in the sense that the simple replacement of AtTic40 IEM targeting signals (S/P-rich domain and TMD) by IEP37 stop-transfer targeting signal (TMD) alone changes the interpretation of the insertion pathway. It corroborates with the data presented in this study that the IEP37 TMD alone is sufficient and necessary to target itself and a mature, otherwise stromal protein (SSU), to the membrane through the stop-transfer pathway. Also in agreement with these data, AtTic40, a post-import substrate, is targeted to the inner envelope membrane via the stop-transfer pathway, when its targeting
signals are replaced by IEP37 TMD. This argues that the TMD of IEP37 functions as a stop-transfer determinant regardless of the position of the TMD along the protein, which in the case of AtTic40-IEP37TM is closer to the N-terminus of the hybrid protein.

Another interesting feature of this construct (pre-AtTic40-IEP37TM) is that it has a very inefficient secondary processing, with the accumulation of the intermediate size form in the membrane fraction. This effect can be due to a few reasons, including the lack of association with the post-import machinery that might recruit the still unknown peptidase responsible for the second processing. Another possibility that was explored and proven to be the case is that due to a different mechanism of insertion, this substrate attains a different membrane topology rendering the processing site inaccessible to the peptidase and resulting in the accumulation of the intermediate sized form. Previous unpublished results in our lab (Li, M., personal communication) have shown that removing the TMD of AtTic40 hindered its IEM insertion, but the stromal localized soluble protein can still be processed to the mature form in the stroma. This is in accordance with the fact that placing the second processing site in the IMS side of the IEM inhibits the second processing, suggesting that the peptidase resides in the stroma.

If we think about the dynamics of protein import in chloroplasts, the transit peptide is recognized by the translocation apparatus and translocation proceeds N-terminus first. Then, the wild-type IEP37, as a stop-transfer substrate with the TMD in the C-terminus, has to flip around and place the bulky N-terminus in the IMS side of the membrane. Upon substitution of IEP37 TMD by AtTic40 TMD (pre-IEP37-Tic40SPTM), the flipping event is inhibited, even with the conservation of the insertion pathway (still a stop-transfer substrate; figures 2.11 and 2.13). This protein is translocated N-terminus
first, and whenever the C-terminus TM helix reaches the translocon, it supposedly functions as a stop-transfer signal prompting the channel to open laterally releasing the protein in the lipid bilayer. Taken together, These results are a demonstration that the IEP37 TMD has dual targeting roles, functioning as a stop-transfer signal and a topology determinant.

Previous reports have not only demonstrated the existence of the post-import pathway using AtTic40 as a substrate, but also, the most recent of them has shown that this protein contains a S/P-rich domain right upstream to the TMD that is essential for membrane insertion (Lubeck et al., 1997; Li and Schnell, 2006; Tripp et al., 2007). Using this information, I analyzed the potential of this signal in directing IEP37 to the IEM through the post-import pathway. Surprisingly, the S/P-rich domain and TMD of AtTic40, which has been shown to be sufficient to target GFP to the IEM via a soluble form (which is an indication of the post-import pathway) was not enough to do the same with IEP37 when fused to its C-terminus. Interestingly, Tripp and colleagues (Tripp et al., 2007) have attempted a similar strategy using the Arc6 protein, in which they replaced the Arc6 TMD by AtTic40 TMD and vice-versa. Neither construct was able to stably integrate in the membrane, which may have been caused by the fact that Arc6 TMD has an inverted orientation in relation to the IEM when compared to AtTic40, thereby hindering the insertion of an inverted TMD in the IEM.

In my assays, I used IEP37 for it has the same membrane topology as AtTic40, making the substitution of TMDs a more reasonable approach. In the case of AtTic40 S/P-rich domain and TMD in the place of IEP37 TMD, the protein was able to reach the membrane and be integrated properly, and even with the presence of the S/P-rich domain,
the post-import pathway was not used (figures 2.10 and 2.11). Therefore, the Tic40 TMD was interpreted by the translocon as a stop-transfer signal, rather than a post-import determinant. This raised a question about the role of the context of this signal during import, for in this construct, the S/P-rich domain and TMD are close to the C-terminus of the protein. On the converse construct, in which Tic40 S/P-rich domain and TMD were replaced by IEP37 TMD, the protein still imported as a stop-transfer substrate, even with the signal closer to the N-terminus (figure 2.7). This phenomenon led me to the interpretation that the stop-transfer pathway seems to be the ‘general’ pathway of membrane protein insertion in the IEM, and the S/P-rich domain might act as a signal to prevent the TMD to stop during translocation, when in the proper context in the protein. This was further corroborated by the fact that placing the Tic40 S/P-rich domain and TMD in the N-terminus of IEP37 targeted this hybrid protein to the IEM through a soluble intermediate, rather than via stop-transfer (figure 2.12). Therefore, it is clear that the post-import pathway relies on a proper positional context of its signals to be interpreted by the translocon as such, and might be achieved by specific context-dependent interactions with the yet unknown machinery that acts in the post-import pathway.

In mitochondria, a similar set of pathways were shown to be present. These two pathways operate through different Tim complexes (Herrmann and Neupert, 2003; Herrmann and Bonnefoy, 2004), and it has been shown that the presence of prolines in the TMD is a fundamental determinant in the capability of a TMD to be arrested or not in the membrane during translocation. It has been demonstrated that prolines in the TMD cause these helices to be transferred by the translocon to the matrix, disfavoring TM
arrest and transfer to the lipid bilayer (Meier et al., 2005) and prompting these proteins to be inserted via a conservative sorting process that is similar to the post-import pathway in chloroplasts. The main difference between these two systems is that a pathway analogous to the conservative sorting was detected in the chloroplast thylakoid membrane, rather than IEM, and is promoted by an insertase Alb3 conserved from the endosymbiotic prokaryote (thereby the name ‘conservative sorting’), and the insertase Oxa1 in the mitochondrial inner membrane.

A closer look at the TM helices of IEP37 and Tic40, the models used in this study for each pathway, reveals that both helices contain a single proline, which by itself, could not be a determinant of the pathway, provided that these proteins adopt different routes. Another evidence for the unimportance of this proline residue in the Tic40 TMD in the transference/arrest of the protein at the translocon was provided by a point mutation of this proline to a leucine which had no effect in the Tic40 insertion pathway, being targeted to the stroma and subsequently redirected to the IEM (Tripp et al., 2007).
In my study, I could determine that the TMD of IEP37 is a stop-transfer signal that operates during chloroplast protein import. It is capable of arresting the protein at the level of the translocon and mediating insertion of the protein into the lipid bilayer without the appearance of a soluble intermediate.

Moreover, the IEP37 TMD also functions as a membrane topology determinant in the native and non-native proteins. Insertion of IEP37 TMD in the membrane occurs only in the N-out/C-in orientation. A similar observation was made previously, in which the transmembrane sequence at the C-terminus of murine surface immunoglobulin heavy chain was shown to have a dual role in triggering the stop-transfer mechanism and also functioning as a topology determinant in the ER (Yost et al., 1983).

Recent studies (Li and Schnell, 2006; Tripp et al., 2007; Chiu and Li, 2008) have demonstrated that the AtTic40 protein is imported into chloroplasts and integrated in the IEM via a soluble intermediate, and a S/P-rich domain adjacent to the TMD is required for membrane insertion. These AtTic40 signals were demonstrated to target GFP to the IEM via a soluble intermediate (Tripp et al., 2007). This prompted me to examine the functionality of these signals in a native IEM protein (IEP37) and analyze the possibility of these signals to divert IEP37 to a post-import pathway. Interestingly, with the addition of AtTic40 TMD alone and even in the presence of the S/P-rich domain, IEP37 was still arrested in the membrane and insertion occurred via the stop-transfer pathway, with no detectable soluble intermediate. This observation raised the possibility that the AtTic40
IEM targeting signals may be context dependent. I then examined this possibility and demonstrated that indeed, if I add the AtTic40 IEM targeting signals to the N-terminus of IEP37 lacking its own TMD, IEP37 is now imported via the post-import pathway. Therefore, I demonstrated that the functionality of the AtTic40 S/P-rich domain and TMD is context dependent.

Therefore, these observations led me to conclude that the stop-transfer pathway is the default pathway for IEM protein insertion because the AtTic40 TMD can function as a stop-transfer signal when placed in the C-terminus of the protein. The S/P-rich domain may, therefore, act by preventing the arrest of the TMD when placed in the N-terminus of AtTic40, but this function cannot be performed at the C-terminus of the protein being threaded in the translocation channel. One possibility to explain the role of the S/P-rich domain is that it may recruit factors that bind to the protein and prevent the TMD from triggering the lateral release at the IEM (stop-transfer integration). These factors could be specific targeting factors that aid post-import insertion or molecular chaperones that maintain the int-AtTic40 in the soluble state when it is released to the stroma.

Moreover, I also examined the topology of the IEP37 constructs harboring the AtTic40 signals and showed that, in contrast to IEP37 TMD, the AtTic40 TMD can insert in the IEM in both directions. The AtTic40 adopts an N-out/C-in orientation in native AtTic40, which is the same topology as IEP37. However, when the Tic40 TMD is placed at the C-terminus of IEP37, this hybrid protein is now in the N-in/C-out orientation in relation to the IEM, as shown by protease treatment of inside-out IEM vesicles.
With these results in hand, I decided to take a closer look at these TMDs in an attempt to detect specific features that would lead the translocon to distinguish between these two pathways.

A slightly higher hydrophobicity of the stop-transfer TMDs have been reported previously for bacterial membrane proteins that use the SecYEG translocase for stop-transfer insertion, in contrast to those TMDs that are inserted with the aid of the YidC insertase (Duong and Wickner, 1998; Xie et al., 2007). Also in bacteria, it has been proposed that the threshold hydrophobicity for a TMD to act as a stop-transfer signal is that equivalent to 16Ala:5Leu (Chen and Kendall, 1995), which corresponds to about 2.0 in the Kyte and Doolittle plot (Kyte and Doolittle, 1982). With these observations, I decided to look at the TMDs of other known chloroplast IEM resident proteins, and found that although the hydrophobicity of the stop-transfer substrates TMD is slightly higher (average 2.0 in the Kyte and Doolittle plot) than the two known post-import substrates (average 1.5) (Table 3.1), the ranges are largely overlapping. Therefore, it is still not clear whether or not it could account for such distinction.

Given the similar endosymbiotic nature of mitochondria and chloroplast, and the numerous studies about its inner membrane protein insertion, I decided to look for common features between these systems. In the case of mitochondria, it’s known that inner membrane proteins can be directly integrated in the membrane by the Tim22 or the Tim23 complex. Moreover, proteins that follow a soluble intermediate traverse the inner membrane through the Tim23 complex (Herrmann and Neupert, 2003).
Table 3.1. Analysis of TMD properties of known IEM resident proteins

| Protein  | Number of predicted TMs | Insertion pathway | Hydrophobicity of TMD (max) (Kyte and Doolittle, 1982) | Distribution of hydrophobicity in the TMD (N to C) | Charge distribution of flanking regions<sup>a</sup> | Helix disruption<sup>b</sup> Prolines <sup>b</sup> |
|----------|-------------------------|-------------------|--------------------------------------------------------|------------------------------------------------|------------------------------------------------|------------------------------------------------|
| IEP37    | 1                       | Stop-transfer<sup>1</sup> | 1.9                                                    | Low-high-low                                        | R-TM-KK                                          | 1C                                                   |
| TPT      | 7                       | Stop-transfer<sup>2</sup> | 1.75                                                   | Variable                                            | KK-TM-KK                                          | -                                                   |
| H17      | 1                       | Stop-transfer<sup>3</sup> | 2.3                                                    | Variable-low                                         | DDD-TM-KDRKD                                      | 1C                                                   |
| HP36     | 9                       | Stop-transfer<sup>4</sup> | 2.1                                                    | Variable-high-variable                               | KEEK-TM-KE                                        | 1                                                   |
| PPT      | 6                       | Stop-transfer<sup>5</sup> | 1.4                                                    | Variable                                            | EDKK-TM-KK                                        | -                                                   |
| AtTic40  | 1                       | Post-import<sup>6</sup> | 1.6                                                    | Low-variable                                         | 0-TM-KK                                          | 1N                                                   |
| AtTic110 | 2                       | Post-import<sup>7</sup> | 1.4                                                    | High–low                                            | EKR-TM-RR                                         | -                                                   |
| AtTic21  | 4                       | Post-import<sup>8</sup> | 2.1                                                    | Low-variable                                         | KRKR-TM-K                                         | -                                                   |

<sup>a</sup>The charged amino acid residues represent those found within the 10 TM helix flanking residues, and their charges.

<sup>b</sup>The N and C represent the closest proximity of the proline residue when found in the TM helix. The absence means that the proline is in the middle of the helix.

<sup>1</sup>(Firlej-Kwoka et al., 2008)
<sup>2</sup>(Knight and Gray, 1995)
<sup>3</sup>(Li and Schnell, 2006; Tripp et al., 2007)
<sup>4</sup>(Lubeck et al., 1997)
<sup>5</sup>(Chiu and Li, 2008)
It has been proposed that the presence of transmembrane helix-disrupting proline residues strongly disfavors the translocation arrest of TMDs by Tim23, and favor the transfer of preproteins to the matrix (Meier et al., 2005). The matrix localized intermediates are then targeted to the inner membrane by the Oxa1 insertase.

I then proceeded to look for the presence of helix disrupting prolines in the TMDs of chloroplast stop-transfer and post-import substrates. According to the data presented in Table 3.1, this feature does not seem to be associated with the discrimination by the Tic complex between transferred or arrested preproteins. Both substrates (e.g. IEP37 and Tic40) have helix disrupting prolines, and the number of prolines is not increased in the case of transferred TMDs, as previously demonstrated for mitochondrial transferred substrates (Meier et al., 2005).

In conclusion, these two models (bacteria and mitochondria) do not point to a feature that can be applied to the chloroplast sorting machinery. Therefore, I looked at other TMD features, in an attempt to find a pattern that would account for this distinction, such as the charged residues distribution in the TMDs flanking regions and the distribution of the hydropobic residues in the TMD (Table 3.1). Unfortunately, none of these characteristics seem to be distinguishable features for the two pathways. Therefore, I conclude that none of these ‘obvious’ characteristics, including sequence identity, are responsible for the pathway distinction at the level of the Tic complex.

However, the number of known substrates for each chloroplast IEM sorting pathways is still very limited, and possibly with the experimental determination of a greater number of substrates and application of more sophisticated bioinformatics tools, the features that may distinguish between these pathways may become more evident.
In a recent study, Firlej-Kwoka and colleagues (Firlej-Kwoka et al., 2008) have determined that several IEM proteins follow the stop-transfer pathway. Nevertheless, a closer look in their published results reveals that three of the analyzed proteins (HP29b, 4TMs; PIC1, 4TMs and PPT, 6TMs) have size intermediates during import. According to their separation method that consists of the simple osmotic lysis, these intermediates were found in the pellet phase and therefore disregarded as potential soluble targeting intermediates. Nevertheless, during my studies, I found that IEP37 lacking its TMD is loosely associated with the membrane and this association can only be disrupted after alkaline extraction. Furthermore, IEP37ΔTM can be re-targeted to the IEM via the post-import pathway when fused to the AtTic40 signals in the right context (Tic40N-IEP37ΔTM). To prove this, I performed an alkaline extraction in a time course experiment and showed that the intermediate is found in the soluble phase and is converted to the mature form. In the case of the proteins analyzed by Firlej-Kwoka and colleagues (Firlej-Kwoka et al., 2008), due to their high hydrophobicity (at least four TMDs), it is valid to speculate that the intermediate-sized protein could be a soluble intermediate that is associated with the membrane, rather than a stop-transfer substrate that is processed twice during translocation, as they suggest. Therefore, a more detailed analysis of these multi-spanning IEM proteins would be a very useful effort towards the understanding of these pathways.

The main question that remains is why the Tic translocon recognizes some TMDs as stop-transfer signals and some not? This applies not only for post-imported substrates, but also for transmembrane proteins that reside in the thylakoid membrane, that are transferred to the stroma before their final sorting to the thylakoid membrane. I therefore
looked at the hydrophobicity of a few transmembrane thylakoidal proteins, but it did not reveal much, for it averaged at about 1.6 in the Kyte and Doolittle scale.

Stop-transfer sequences from ER proteins are also unable to arrest the translocation at the chloroplast envelope level (Lubben et al., 1987).

Therefore, it seems clear that the properties of the stop-transfer signals differ considerably between the chloroplast IEM and the bacterial export, mitochondria and even thylakoid membrane proteins. These observations argue that the translocation machinery at the chloroplast IEM differs significantly from all these systems, and the signals to promote membrane insertion may not be conserved across these translocators.

In fact, the details that underline the functionality of these targeting signals and the channel specificities are still poorly understood. It would be interesting to use the same kind of approach that I used in this study with TMDs of IEM resident proteins that are stop-transfer substrates and TMDs of thylakoid membrane proteins. For instance, it would be particularly informative to address whether or not a thylakoid stop-transfer signal such as the one on the C-terminus of the cytochrome $f$ (Mould et al., 1997) could promote IEM integration by arresting the protein at the level of the IEM. Chloroplast genome expression of an IEP37 mutant containing a thylakoid lumen signal sequence would also be an interesting assay to determine if its TMD is able to be interpreted as such by the thylakoid SecA-translocon.

A few studies (Inaba et al., 2005; Chiu and Li, 2008) have suggested the involvement of the Tic complex in the insertion of post-import substrates. Although the details of the post-import pathway are not fully understood, it can be assumed that the Tic complex mediate both pathways.
Unlike mitochondria, where there is a insertase (Oxa1) in the inner membrane, the chloroplast IEM seems not to have such insertase. This insertase homolog is, in fact, found at the thylakoid membrane (YidC). Proteomic studies have failed to find any proteins with insertase properties (Ferro et al., 2003; Rolland et al., 2003) in the chloroplast IEM. This does not exclude the possibility that a novel mechanism might have evolved in chloroplast to target these proteins to the IEM. However, it does suggest that the players involved in these two pathways may not in fact be entirely distinct, and in fact, these pathways share common elements such as AtTic110, AtTic40 and Hsp93. Therefore, it is possible that the Tic complex is responsible for the sorting of these two classes of proteins, as well as for the IEM re-insertion of proteins that follow the post-import pathway.
5.1 Generation of constructs

I obtained the vectors containing wt IEP37 (locus AT3G63410), AtTic40 (locus At5g16620) and pAts1B (SSU – Rubisco small subunit) (locus At5g38430) from the Arabidopsis Biological Resource Center (ABRC). All cDNAs were cloned into the pET21d vector (Novagen) under control of the T7 promoter sequence. These constructs were used as base for all the deletions and swaps that followed.

For all the overlap extension PCRs (OE-PCRs), I followed the standard protocol (Horton et al., 1989; Horton et al., 1993) with modifications as follows. All reactions were performed with the same set of conditions for all constructs, optimizing the annealing temperature and extension time depending on the sequence to be overlapped and the set of primers being used. Briefly, I prepared the reactions without primers and used an initial denaturing step of 5 minutes at 95°C, followed by 10 cycles of a 1-minute denaturing step at 95°C, 30-second annealing step at 50°C and 30-90-second extension at 72°C. After the 10 cycles, there was a final elongation step at 72°C for 7 minutes. After these initial 10 cycles, the primers were added and the reaction was repeated with the same conditions for an additional 32 cycles. Table 4.1 at the end of this chapter shows the list of primers used for each sub-step of the OE-PCRs, along with their nucleotide sequence and the OE-PCRs steps to obtain each construct.
5.2 Chloroplast isolation and import assays

The isolation of intact chloroplasts to be used in the in vitro import assays was performed as described by (Smith et al., 2003). All the procedures for chloroplast isolation were carried out in a cold room, all the solutions were chilled on ice and the chloroplasts always kept on ice and used in the same day. Briefly, the leaf tissue was harvested from plants 8-12-day-old pea seedlings grown on soil and ground in a rotary homogenizer in chilled grinding buffer (25 mM Hepes-KOH pH 8, 330 mM sorbitol, 2 mM EDTA, 1 mM MgCl$_2$, 1 mM MnCl$_2$, 0.25% w/v BSA, 0.1% w/v sodium ascorbate, 0.05% v/v protease inhibitor cocktail (SIGMA P-9599)). The homogenate was passed through three layers of Miracloth and centrifuged for 2 min at 1500 x g, 4 °C. Most of the supernatant was discarded and the approximately 3-5 ml of the remainder was used to resuspend the pellet by gently swirling. The resuspension was then transferred to the top of a percoll gradient (7 mL of 45% percoll over 5 mL of 85% percoll, 25 mM Hepes-KOH pH 8, 330 mM sorbitol, 2 mM EDTA, 1 mM MgCl$_2$, 0.1% w/v sodium ascorbate). The gradient was centrifuged for 10 min at 1500 x g, 4 °C with the centrifuge brakes off. The 85%:40% percoll interface containing the intact chloroplasts was collected with a Pasteur pipette, diluted in HS buffer (25 mM Hepes-KOH pH 8, 330 mM sorbitol) and centrifuged for 2 min at 4500 x g, 4 °C. The HS wash was repeated once and the pelleted chloroplasts were resuspended in HS, and an aliquot was used to measure the 652nm absorbance (in 80% acetone) to determine the chlorophyll concentration, which was adjusted to 1-2 mg/ml.

For the import assays, I used in vitro translated protein (IVT) generated with the Promega TNT® T7 Coupled Reticulocyte Lysate System (Cat.# L4610), according to the
manufacturer’s instructions. The reaction mix (import buffer) contained ~500 µg of chlorophyll (isolated chloroplasts), 5mM Mg(CH₃COO)₂, 25 mM CH₃COOK, 5mM ATP, 1mM DTT, 1mM methionine, in HS buffer. It was pre-incubated at 26 °C for 10 minutes before the addition of 25 µL of IVT, and then import proceeded at the same temperature for the times indicated in each figure. The reactions were stopped by adding 2 volumes of ice-cold HS buffer, placing on ice and immediately centrifuging at 900 x g for 3 minutes. Protease treatment of the chloroplast followed when indicated using 200 µg/mL of thermolysin, on ice for 30 min. Thermolysin was inactivated by addition of EDTA to a final concentration of 20 mM. After import (and protease treatment of indicated samples), the intact chloroplasts were re-isolated through a 40% percoll cushion, a fraction taken for quantification, and then processed either for fractionation or directly to SDS-PAGE (resuspension in gel loading buffer).

5.3 Chloroplast lysis and fractionation

The samples that were further separated into membrane and soluble fractions followed either fractionation method:

5.3.1. Osmotic lysis

An excess of 10 volumes of 2 mM EDTA was added to the reaction, which was vortexed and left on ice for 10 minutes, followed by the addition of NaCl to the concentration of 250 mM and centrifuged at 14,000 x g for 30 minutes at 4 °C.
5.3.2. Alkaline extraction

An excess of 10 volumes of 0.2 M Na$_2$CO$_3$ pH 11.5 was added to the reaction, which was vortexed, homogenized with 10 strokes with a Dounce homogenizer and centrifuged at 14,000 x g for 30 minutes at 4 °C.

For both fractionation methods, the pellet was directly processed for SDS-PAGE, whereas the supernatant was mixed with TCA to a final concentration of 20%, which after vortexing was incubated overnight at 4 °C. The proteins were then precipitated by centrifugation at 14,000 x g for 30 minutes at 4 °C, followed by a wash with 0.5% TCA and centrifugation at the same conditions for 15 minutes. The pellet was then re-suspended in gel loading buffer.

5.4 Chase experiments

The chloroplast import reaction was prepared as described in 5.2, but the pre-incubation was at 20°C and the import proceeded for only 5 minutes at the same temperature. These conditions were shown in our lab to enrich the soluble intermediate form of AtTic40 (data not shown). The reaction was immediately stopped and after thermolysin digestion and re-isolation, the chloroplasts were re-introduced in pre-heated (26 °C) import buffer and the reaction proceeded at 26 °C for 0, 5 and 60 minutes. After each time point, the reaction was stopped and fractionation proceeded as described in section 5.3.
5.5 Chloroplast inner-envelope membrane isolation and proteolytic treatment

The isolation of IEM vesicles was performed as described by (Smith et al., 2003) with modifications described as follows. In order to isolate IEMs, after import, the samples were treated with thermolysin as indicated in section 5.2, re-isolated and resuspended in 0.6 M sucrose/TE/DTT buffer (50 mM tricine pH 7.5, 2 mM EDTA, 1 mM DTT). The chloroplasts were frozen at –20 °C overnight. Then, after thawing on ice, ~1-2 mg of pre-frozen isolated chloroplasts (in 0.6 M sucrose/TE/DTT buffer) were added to the imported chloroplasts, the concentration of sucrose adjusted to 0.2 M by addition of TE/DTT buffer and homogenized by 20 strokes at a Dounce homogenizer. The homogenate was then centrifuged at 40,000 x g for 1 hour at 4 °C, and the pellet containing the membranes resuspended in 0.2M sucrose/TE/DTT and homogenized by 20 strokes at a Dounce homogenizer. The homogenate was then loaded onto the top of a 1.0M:0.8M:0.46M sucrose/TE/DTT step gradient and centrifuged 40,000 x g for 1.5 hours at 4 °C. The 1.0M:0.8M interface enriched in IEM vesicles was collected, diluted in 5-10 times HS buffer and centrifuged 40,000 x g for 1.5 hours at 4 °C. The pellets where then resuspended in HS to a concentration of ~ 1µg/ml total protein, measured by the BCA assay.

The vesicles were further treated with thermolysin for 30 minutes on ice in the concentrations described in the figures, with 10 µg of total protein loaded in each gel lane. In the case of the protease activity controls, Triton X-100 was added to the protease assay to a final concentration of 2%. After protease treatment, the reaction was inhibited by the addition of EDTA to a concentration of 20 mM and processed for SDS-PAGE.
Table 4.1 List and nucleotide sequence of primers used in deletion, hybrid and substitution mutants, and reaction steps to obtain each construct.

| Construct          | Reaction n. | Template | Primer set                        | Sequence                                                                 |
|--------------------|-------------|----------|-----------------------------------|--------------------------------------------------------------------------|
| IEP37∆TM           | 1.1         | IEP37    | IEP37-Nco-F (atIEP37-5′NcoI)      | 5′TTAAAACCATGGCTCTTTGATGC TCAACGGG                                      |
|                    |             |          | IEP37∆TM-Not-R                    | 5′CAAGGCGGCGCTCATCCAAAGA AGGAGAA                                       |
| IEP37-Tic40TM      | 2.1         | Tic40    | Tic40TM-IEP37-F                   | 5′AACCCCTTTCTCTTTGAGGAGGA TCACCACCTTTCTGG                               |
|                    |             |          | Tic40TM-IEP37-R                   | 5′TCAGAGGTTTGTTTTGAGGAGGA CGATCTGATCTTTTGTAAATTTGA AGT                |
|                    | 2.2         | IEP37    | IEP37-Nco-F                       | See above                                                               |
|                    |             |          | IEP37∆TM-R                        | 5′TCCCAAGAAGGAGAAGGG                                                   |
|                    | 2.3         | Products from reactions 2.1 and 2.2 | IEP37-Nco-F                       | See above                                                               |
|                    |             |          | IEP37-BamHI-R (atIEP37-3′BamHI)   | 5′TAAGGATCCCTCAGGTTTGGTGTGTTTGGG AGGAGAAGG                               |
| Tic40-IEP37TM      | 3.1         | IEP37    | IEP37TM-Tic40-F                   | 5′CCACCATCTTCTCAACCATAACGC TTCTCTTTGGGATACT                               |
|                    |             |          | IEP37TM-Tic40-R                   | 5′CATAGCTGTGTTGCTGATCCACTTGTAGAT                                       |
|                    | 3.2         | Tic40    | Tic40-Nco-F (AtTic40-5′NcoI)      | 5′TTAAAACCATGGAGAAACCTTACC CTA                                         |
|                    |             |          | Tic40ΔTM-R                        | 5′TATGGTTGATGAAGATGG                                                   |
|                    | 3.3         | Tic40    | Tic40ΔTM-F                        | 5′TATGCAATGCAAACAGCT                                                  |
|                    |             |          | Tic40-BamHI-Rev                   | 5′CCCAGGAAATGACGGGTTGGAGGAT CCTTT                                     |
|                    | 3.4         | Products from reactions 3.1 and 3.2 | Tic40-Nco-F                       | See above                                                               |
|                    |             |          | IEP37TM-Tic40-R                   | See above                                                               |
|                    | 3.5         | Products from reactions 3.3 and 3.4 | Tic40-Nco-F                       | See above                                                               |
|                    |             |          | Tic40-Bam-R                       | See above                                                               |
| Tic40SP-IEP37TM | 4.1 | IEP37 | IEP37-Nco-F | See above |
|-----------------|-----|-------|--------------|-----------|
|                 |     |       | IEP37-ATM-R | See above |
| 4.2             | Tic40 | Tic40SP-IEP37-ATM-F | 5’AACCCCTTCTCTTTTTGACGT | See above |
|                 |     | Tic40TM-IEP37-R | GATCAACAGACAACCTTC | |
| 4.3             | Products from reaction 4.1 and 4.2 | IEP37-Nco-F | See above |
|                 |     | Tic40TM-IEP37-R | See above |
| 4.4             | Product from reaction 4.3 | IEP37-Nco-F | See above |
|                 |     | IEP37-BamHI-R | See above |
| Tic40ΔSP-IEP37TM | 5.1 | Tic40 | Tic40-Nco-F | See above |
|                 |     | Tic40-TP-R | 5’ACGACTCGAAGAAAAATATAC | |
| 5.2             | Tic40 | Tic40-ATM-F | 5’TATGCAATGCAAACAGC | See above |
|                 |     | Tic40-BamHI-R | | |
| 5.3             | IEP37 | Tic40ΔSPATM-IEP37-288-328-Fa | 5’CAGAGGAAGCAAGGCTTTTGCAGG | See above |
|                 |     | Tic40TM-Tic40-R | TCCAAGGAAGGACGTTAGAG | |
| 5.4             | Products from reaction 5.1 and 5.3 | IEP37-Nco-F | See above |
|                 |     | Tic40TM-Tic40-R | See above |
| 5.5             | Products from reaction 5.4 and 5.2 | IEP37-Nco-F | See above |
|                 |     | Tic40-BamHI-R | See above |
| Tic40N-IEP37△TM | 6.1 | Tic40 | Tic40-Nco-F | See above |
|                 |     | Tic40-TP-R | 5’TTCCTCTTAAATTTGAGTTAC | |
| 6.2             | IEP37 | Tic40TM-IEP37-ATP-F | 5’ACTTCAAATTTAAGAAAATCGTC | See above |
|                 |     | Tic40TM-IEP37-Not-R | TTCCCGCCCATCGGC | |
| 6.3             | Products from reaction 6.1 and 6.2 | IEP37-Nco-F | See above |
|                 |     | IEP37△TM-Not-R | See above |
| IEP37-mSSU-IEP37TM | 7.1 | IEP37 | IEP37-Nco-F | See above |
|------------------|-----|-------|-------------|-----------|
|                  |     |       | IEP37-TP-R | 5’CACGCTGCTGCTGC |
| 7.2              | pAts1B (pSSU) | IEP3P-mSSU-F | 5’GCAGCAGCAGCAGCGTGTCATG AAGGTAAGGCC |
|                  | SSU∆Stop-R | 5’AGCATCAGTGAAGCTTGG |
| 7.3              | IEP37 | SSU-IEPCter-F | 5’CCAAGCTTTCTAGATGCAGTCC AAAGGAAGAGG |
|                  |       | IEP37-Bam-R | See above |
| 7.4              | Products from reactions 7.1, 7.2 and 7.3 | IEP37-Nco-F | See above |
|                  |       | IEP37-Bam-R | See above |

Note 1: unless otherwise noted, all vector templates were a 1-2 ng of the cDNA of the described gene cloned in the pET21d vector (Novagen). All the reactions were performed in a final volume of 20µl, and 1µl of each product was used in the OE-PCR, when used as templates.
Alberts, B., Johnson, A, Lewis, J, Raff, M, Roberts, K, Walter, P. (2002). Molecular Biology of the Cell. (Garland Science).

Archibald, J.M. (2009). The puzzle of plastid evolution. Current biology : CB 19, R81-88.

Aronsson, H., and Jarvis, P. (2009). The Chloroplast Protein Import Apparatus, Its Components, and Their Roles. In Plant Cell Monographs - The Chloroplast (Springer Berlin / Heidelberg), pp. 89-123.

Becker, T., Hritz, J., Vogel, M., Caliebe, A., Bukau, B., Soll, J., and Schleiff, E. (2004). Toc12, a novel subunit of the intermembrane space preprotein translocon of chloroplasts. Molecular biology of the cell 15, 5130-5144.

Blobel, G., Walter, P., Chang, C.N., Goldman, B.M., Erickson, A.H., and Lingappa, V.R. (1979). Translocation of proteins across membranes: the signal hypothesis and beyond. Symp Soc Exp Biol 33, 9-36.

Block, M.A., Joyard, J., and Douce, R. (1991). Purification and characterization of E37, a major chloroplast envelope protein. FEBS letters 287, 167-170.

Bolter, B., Soll, J., Schulz, A., Hinnah, S., and Wagner, R. (1998). Origin of a chloroplast protein importer. Proceedings of the National Academy of Sciences of the United States of America 95, 15831-15836.

Bruce, B.D. (2000). Chloroplast transit peptides: structure, function and evolution. Trends in cell biology 10, 440-447.

Chen, H., and Kendall, D.A. (1995). Artificial transmembrane segments. Requirements for stop transfer and polypeptide orientation. The Journal of biological chemistry 270, 14115-14122.

Chiu, C.C., and Li, H.M. (2008). Tic40 is important for reinsertion of proteins from the chloroplast stroma into the inner membrane. The Plant journal : for cell and molecular biology 56, 793-801.

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. The EMBO journal 22, 2970-2980.

Cline, K., and Henry, R. (1996). Import and routing of nucleus-encoded chloroplast proteins. Annual review of cell and developmental biology 12, 1-26.

Cline, K., and Dabney-Smith, C. (2008). Plastid protein import and sorting: different paths to the same compartments. Current opinion in plant biology 11, 585-592.

Dalbey, R.E., and Robinson, C. (1999). Protein translocation into and across the bacterial plasma membrane and the plant thylakoid membrane. Trends Biochem Sci 24, 17-22.

Dalbey, R.E., and Kuhn, A. (2000). Evolutionarily related insertion pathways of bacterial, mitochondrial, and thylakoid membrane proteins. Annual review of cell and developmental biology 16, 51-87.
Di Cola, A., Klostermann, E., and Robinson, C. (2005). The complexity of pathways for protein import into thylakoids: it's not easy being green. Biochemical Society transactions 33, 1024-1027.

Dreses-Werringloer, U., Fischer, K., Wachter, E., Link, T.A., and Flugge, U.I. (1991). cDNA sequence and deduced amino acid sequence of the precursor of the 37-kDa inner envelope membrane polypeptide from spinach chloroplasts. Its transit peptide contains an amphiphilic alpha-helix as the only detectable structural element. European journal of biochemistry / FEBS 195, 361-368.

Duong, F., and Wickner, W. (1998). Sec-dependent membrane protein biogenesis: SecYEG, preprotein hydrophobicity and translocation kinetics control the stop-transfer function. The EMBO journal 17, 696-705.

Ferro, M., Salvi, D., Brugiere, S., Miras, S., Kowalski, S., Louwagie, M., Garin, J., Joyard, J., and Rolland, N. (2003). Proteomics of the chloroplast envelope membranes from Arabidopsis thaliana. Molecular & cellular proteomics : MCP 2, 325-345.

Firlej-Kwoka, E., Strittmatter, P., Soll, J., and Bolter, B. (2008). Import of preproteins into the chloroplast inner envelope membrane. Plant molecular biology 68, 505-519.

Flugge, U.I., Fischer, K., Gross, A., Sebald, W., Lottspeich, F., and Eckerskorn, C. (1989). The triose phosphate-3-phosphoglycerate-phosphate translocator from spinach chloroplasts: nucleotide sequence of a full-length cDNA clone and import of the in vitro synthesized precursor protein into chloroplasts. The EMBO journal 8, 39-46.

Gould, S.B., Waller, R.F. and McFadden, G.I. (2008). Plastid Evolution. Annu Rev Plant Biol 59, 491-517.

Herrmann, J.M., and Neupert, W. (2003). Protein insertion into the inner membrane of mitochondria. IUBMB life 55, 219-225.

Hinnah, S.C., Hill, K., Wagner, R., Schlicher, T., and Soll, J. (1997). Reconstitution of a chloroplast protein import channel. The EMBO journal 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. Biophysical journal 83, 899-911.

Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K., and Pease, L.R. (1989). Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene 77, 61-68.

Horton, R.M., Ho, S.N., Pullen, J.K., Hunt, H.D., Cai, Z., and Pease, L.R. (1993). Gene splicing by overlap extension. Methods in enzymology 217, 270-279.

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. The Journal of biological chemistry 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. The Plant cell 17, 1482-1496.

Jackson-Constan, D., and Keegstra, K. (2001). Arabidopsis genes encoding components of the chloroplastic protein import apparatus. Plant physiology 125, 1567-1576.
Jarvis, P. (2008). Targeting of nucleus-encoded proteins to chloroplasts in plants. The New phytologist 179, 257-285.

Jarvis, P., and Robinson, C. (2004). Mechanisms of protein import and routing in chloroplasts. Current biology : CB 14, R1064-1077.

Joyard, J., Teyssier, E., Miege, C., Berny-Seigneurn, 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 physiology 118, 715-723.

Keegstra, K., and Yousif, A.E. (1986). Isolation and characterization of chloroplast envelope membranes. Methods in enzymology 118, 316-325.

Kessler, F., and Schnell, D.J. (2006). The function and diversity of plastid protein import pathways: a multilane GTPase highway into plastids. Traffic (Copenhagen, Denmark) 7, 248-257.

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. The Plant cell.

Knight, J.S., and Gray, J.C. (1995). The N-terminal hydrophobic region of the mature phosphate translocator is sufficient for targeting to the chloroplast inner envelope membrane. The Plant cell 7, 1421-1432.

Ko, K., and Cashmore, A.R. (1989). Targeting of proteins to the thylakoid lumen by the bipartite transit peptide of the 33 kd oxygen-evolving protein. The EMBO journal 8, 3187-3194.

Kobayashi, K., Nakamura, Y., and Ohta, H. (2009). Type A and type B monogalactosyldiacylglycerol synthases are spatially and functionally separated in the plastids of higher plants. Plant Physiol Biochem 47, 518-525.

Kouranov, A., and Schnell, D.J. (1997). Analysis of the interactions of preproteins with the import machinery over the course of protein import into chloroplasts. The Journal of cell biology 139, 1677-1685.

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. The Journal of cell biology 143, 991-1002.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105-132.

Leister, D. (2003). Chloroplast research in the genomic age. Trends in genetics : TIG 19, 47-56.

Li, M., and Schnell, D.J. (2006). Reconstitution of protein targeting to the inner envelope membrane of chloroplasts. The Journal of cell biology 175, 249-259.

Lubben, T.H., Bansberg, J., and Keegstra, K. (1987). Stop-Transfer Regions Do Not Halt Translocation of Proteins into Chloroplasts. Science (New York, N.Y.) 238, 1112-1114.

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

Lynch, M., and Blanchard, J.L. (1998). Deleterious mutation accumulation in organelle genomes. Genetica 102-103, 29-39.

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. The Journal of cell biology 134, 315-327.

Martin, W., and Herrmann, R.G. (1998). Gene transfer from organelles to the nucleus: how much, what happens, and Why? Plant physiology 118, 9-17.

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

Meier, S., Neupert, W., and Herrmann, J.M. (2005). Proline residues of transmembrane domains determine the sorting of inner membrane proteins in mitochondria. The Journal of cell biology 170, 881-888.

Motohashi, R., Ito, T., Kobayashi, M., Tajii, T., Nagata, N., Asami, T., Yoshida, S., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2003). Functional analysis of the 37 kDa inner envelope membrane polypeptide in chloroplast biogenesis using a Ds-tagged Arabidopsis pale-green mutant. The Plant journal : for cell and molecular biology 34, 719-731.

Mould, R.M., Knight, J.S., Bogsch, E., and Gray, J.C. (1997). Azide-sensitive thylakoid membrane insertion of chimeric cytochrome f polypeptides imported by isolated pea chloroplasts. The Plant journal : for cell and molecular biology 11, 1051-1058.

Rapoport, T.A. (1986). Protein translocation across and integration into membranes. CRC Crit Rev Biochem 20, 73-137.

Reyes-Prieto, A., Weber, A.P., and Bhattacharya, D. (2007). The origin and establishment of the plastid in algae and plants. Annual review of genetics 41, 147-168.

Robinson, C., and Bolhuis, A. (2004). Tat-dependent protein targeting in prokaryotes and chloroplasts. Biochimica et biophysica acta 1694, 135-147.

Rolland, N., Ferro, M., Seigneurin-Berny, D., Garin, J., Douce, R., and Joyard, J. (2003). Proteomics of chloroplast envelope membranes. Photosynthesis research 78, 205-230.

Sagan, L. (1967). On the origin of mitosing cells. Journal of theoretical biology 14, 255-274.

Sargent, F. (2007). The twin-arginine transport system: moving folded proteins across membranes. Biochemical Society transactions 35, 835-847.

Schnell, D.J. (1998). Protein targeting to the thylakoid membrane. Annu. Rev. Plant Physiol. Plant Mol. Biol. 49, 97-126.

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. The Journal of cell biology 120, 103-115.

Schnell, D.J., Kessler, F., and Blobel, G. (1994). Isolation of components of the chloroplast protein import machinery. Science (New York, N.Y.) 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 in cell biology 7, 303-304.

Schwacke R, S.A., Van Der Graaff E, Fischer K, Catoni E, Desimone M, Frommer WB, Flugge UI, Kunze R. (2003). ARAMEMNON, a Novel Database for Arabidopsis Integral Membrane Proteins. Plant physiology 131, 16-26.
Singh, N.D., Li, M., Lee, S.B., Schnell, D., and Daniell, H. (2008). Arabidopsis Tic40 expression in tobacco chloroplasts results in massive proliferation of the inner envelope membrane and upregulation of associated proteins. The Plant cell 20, 3405-3417.

Smeekens, S., Bauerle, C., Hageman, J., Keegstra, K., and Weisbeek, P. (1986). The role of the transit peptide in the routing of precursors toward different chloroplast compartments. Cell 46, 365-375.

Smith, M.D., Schnell, D.J., Fitzpatrick, L., and Keegstra, K. (2003). In vitro analysis of chloroplast protein import. Current protocols in cell biology / editorial board, Juan S. Bonifacino ... [et al.] Chapter 11, Unit11 16.

Soll, J., Schultz, G., Joyard, J., Douce, R., and Block, M.A. (1985). Localization and synthesis of prenylquinones in isolated outer and inner envelope membranes from spinach chloroplasts. Archives of biochemistry and biophysics 238, 290-299.

Stahl, T., Glockmann, C., Soll, J., and Heins, L. (1999). Tic40, a new "old" subunit of the chloroplast protein import translocon. The Journal of biological chemistry 274, 37467-37472.

Timmis, J.N., Ayliffe, M.A., Huang, C.Y., and Martin, W. (2004). Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nature reviews. Genetics 5, 123-135.

Tripp, J., Inoue, K., Keegstra, K., and Froehlich, J.E. (2007). A novel serine/proline-rich domain in combination with a transmembrane domain is required for the insertion of AtTic40 into the inner envelope membrane of chloroplasts. The Plant journal : for cell and molecular biology 52, 824-838.

van Loon, A.P.G.M., and Schatz, G. (1987). Transport of proteins to the inner membrane: the "sorting" domain of the cytochrome c1 presequence is a stop-transfer sequence for the mitochondrial inner membrane. EMBO J. 6, 2441-2448.

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

von Heijne, G., and Nishikawa, K. (1991). Chloroplast transit peptides. The perfect random coil? FEBS letters 278, 1-3.

Waegemann, K., Paulsen, H., Soll, J. (1990). Translocation of proteins into isolated chloroplasts requires cytosolic factors to obtain import competence. FEBS letters 261, 89-92.

Weber, A.P., Schwacke, R., and Flugge, U.I. (2005). Solute transporters of the plastid envelope membrane. Annu Rev Plant Biol 56, 133-164.

Xie, K., Hessa, T., Seppala, S., Rapp, M., von Heijne, G., and Dalbey, R.E. (2007). Features of transmembrane segments that promote the lateral release from the translocase into the lipid phase. Biochemistry 46, 15153-15161.

Yost, C.S., Hedgpeth, J., and Lingappa, V.R. (1983). A stop transfer sequence confers predictable transmembrane orientation to a previously secreted protein in cell-free systems. Cell 34, 759-766.