國立臺灣大學分子與細胞生物學研究所

碩士論文

Institute of Molecular and Cellular Biology
National Taiwan University

Master Thesis

葉綠體內膜運輸機組蛋白 Tic40 之導引序列分析

Functional Characterizations of Sequence Motifs in the Transit Peptide of *Pisum sativum* Tic40

詹博婷

PO-TING CHAN

指導教授：李秀敏 博士

Advisor: Hsou-min Li, Ph.D.

中華民國 100 年 6 月

June, 2011

目錄

Abstract

Most chloroplast proteins are encoded in the nucleus, translated in the cytosol, and then transported into chloroplasts through the Toc (Translocon at the outer envelope membrane of chloroplasts) and Tic (Translocon at the inner envelope membrane of chloroplasts) complexes. Nucleus-encoded chloroplast proteins are usually synthesized as higher molecular weight precursors with N-terminal targeting signals called the transit peptides. Even though transit peptides are necessary and sufficient for chloroplast precursor proteins import, what constitute the critical sequence features for receptor recognition in the transit peptides is still poorly understood. Previous research has shown that precursors can be classified into three groups based on their preference to be imported into chloroplasts of different ages and the age-selective signal is located within the transit peptide of each precursor. In this study, I tried to identify the transit peptide features that confer preference for older chloroplasts. I divided the transit peptide of prPsTic40 (precursor of Tic40 from pea *Pisum sativum*), one of the precursors that prefer to be imported into older chloroplasts, into eight blocks and generated serial alanine scanning mutants. When block 4 (residues 28-36) was mutated, import efficiency into older chloroplasts was significantly decreased, while import efficiency into younger chloroplasts was not affected. Further mutations within block 4 through site-directed mutagenesis revealed that two positive charges, arginine and lysine of residues 29 and 30, are necessary and sufficient to confer preference for older chloroplasts. Moreover, two positive charges in the transit peptide of prAtL11 (precursor of L11 from *Arabidopsis thaliana*), another precursor that also prefers to be imported into older chloroplasts, are also responsible for prAtL11 preference for older chloroplasts. These results indicate that two consecutive positive charges in the transit peptide are the critical feature constituting the age-selective signal for older chloroplasts preference.

中文摘要

葉綠體是植物細胞所具有的特殊胞器，其功能主要為進行光合作用以提供養份使植物生長。大部分的葉綠體蛋白質都是藉由細胞核內基因所轉錄，於細胞質中轉譯，再自細胞質送入葉綠體執行功能。在還未送入葉綠體前，細胞質內的葉綠體前驅蛋白質(precursor protein)在 N 端帶有一段可被葉綠體膜上受器所辨認的導引序列(transit peptide)，之後便藉由葉綠體外膜及內膜上的運輸機組(translocon complex)送入葉綠體。而該段導引序列在前驅蛋白質被送入葉綠體後，即會被葉綠體基質處理酶(stromal processing peptidase)所切除而成為成熟蛋白質(mature protein)。目前的研究成果已經確認葉綠體蛋白質之導引序列，對於葉綠體蛋白質的運輸是扮演著一個充分且必要的角色。但這些導引序列無論是胺基酸組成或是長度，都存在著高度相異性，因此關於葉綠體蛋白質的導引序列，至今我們仍然不甚瞭解。

過去關於葉綠體蛋白質的研究發現：葉綠體蛋白質可依照其喜歡的葉綠體年紀分為三大類，分別是喜歡進年輕葉綠體、喜歡進年老葉綠體、以及在年輕或年老葉綠體皆無差異者；而這種對年齡選擇的能力，是存在於葉綠體的導引序列。在這篇論文中，我主要針對喜歡進年老葉綠體的蛋白質進行導引序列功能性分析。利用丙氨酸置換(alanine scanning)及定點突變(site-directed mutagenesis)等技術將豌豆 Tic40 前驅蛋白質(prPsTic40)之導引序列進行區段分析，發現影響偏好進年老葉綠體的關鍵區域位在皆帶正電的第 29 個精氨酸(arginine)與第 30 個賴氨酸(lysine)。當拿掉這兩個胺基酸時，就會發現 Tic40 不喜歡進老的葉綠體；而再把這兩個胺基酸放回導引序列後，便可發現 Tic40 恢復喜歡進年老葉綠體的特性。除此之外，在另一個也具有喜歡進年老葉綠體特性的蛋白質 prAtL11，其導引序列也發現兩個帶正電的 lysine 扮演著同樣會影響年紀選擇的角色。這些研究結果顯示：偏好進年老葉綠體的蛋白質，其位於導引序列的連續兩個攜帶正電氨基酸，是影響年紀選擇的關鍵。

Introduction

Chloroplasts are intracellular organelles that are the site of photosynthesis in plants. The origin of chloroplasts is thought to have occurred 1 to 2 billion years ago through endosymbiosis. During the endosymbiotic process, a photosynthetic cyanobacterium was taken up by a heterotrophic protist and then eventually transformed into a double-membrane-bound organelle within the host cell [1].

During evolution, most genes from the original cyanobacterial endosymbion have been transferred to the host nucleus [2, 3]. Thus, proteins encoded by these genes have to be imported back into chloroplasts to keep chloroplasts functional [4]. Machinery on the chloroplast envelope, the translocon, is responsible for translocating nucleus-encoded chloroplast proteins across the double-membrane envelope of chloroplasts . Most nucleus-encoded chloroplast proteins are synthesized as higher molecular weight precursors in cytosol with cleavable N-terminal targeting sequences called transit peptides, which are recognized by the translocon during import into chloroplasts.

Translocon components located in the outer envelope membrane are called Toc (**t**ranslocon at the **o**uter-envelope membrane of **c**hloroplasts) proteins and those in the inner envelope membrane are called Tic (**t**ranslocon at the **i**nner-envelope membrane of **c**hloroplasts) proteins. Three Toc proteins, Toc159, Toc75 and Toc34, form a stable Toc complex and constitute the major pathway for import across the outer membrane. Toc159 and Toc34 mediate precursor protein recognition. In higher plants, Toc159 is encoded by a multigene family and different family members have different precursor preferences [5, 6]. Toc75 is the major channel constituent for precursor protein translocation across the outer membrane [7, 8].

After precursor proteins are translocated through the Toc complex, they require Tic proteins to translocate them across or into the inner membrane. Tic proteins required for inner membrane translocation include Tic20, Tic21, Tic40 and Tic110. Tic20 and Tic21 are channels across the chloroplast inner membrane. Both proteins are essential for chloroplast biogenesis; the former is important for early stages and the latter is more important for mature stages of chloroplast biogenesis. The stromal domain of Tic110 contains a transit peptide-binding site, which is the initial site for precursor proteins binding while they emerge from the Tic complex. Tic110 also functions as a scaffold for the recruitment of stromal chaperones and cochaperones, including Hsp93 and Tic40, to import sites during inner membrane translocation. Hsp93 is likely a motor that provides the driving force for chloroplast protein import into stoma. Tic40 with a large hydrophilic domain in the stroma functions as a membrane-anchored cochaperone. Tic40 stimulates transit peptide release from Tic110 and then ATP hydrolysis by Hsp93, thus coordinates the actions of Tic110 and Hsp93. After translocation across the inner membrane, transit peptides are removed by a zinc-binding metallopeptidase, the stromal processing peptidase (SPP). Ultimately, the transit peptide is converted to a degradable subfragment and the processed precursor is converted to the mature form of a functional chloroplast protein [7, 9].

How targeting signals are recognized by receptors is best studied for protein import into the endoplasmic reticulum (ER) and mitochondria [10]. Sequence analyses have indicated that the ER targeting signal peptide contains one or more positively charged amino acids followed by a continuous stretch of 6-12 hydrophobic residues. Deletion of several amino acids or the insertion of charged amino acids from the hydrophobic core can abolish the targeting ability of the ER-directed signal peptide; thus the hydrophobic core of ER signal peptides is essential for ER

precursors' translocation [11].

ER precursor protein targeting is mediated by the signal recognition particle (SRP) and its membrane-localized receptor. The translocation of ER precursors is initiated when the most N-terminal 16- to 30- residues of the signal peptide in nascent proteins emerge from the ribosome and are bound by the SRP. Cross-linking experiments have shown that the signal peptides of ER precursor proteins are recognized by the 54-kD subunit of the SRP (SRP54). SRP54 is composed of three domains: a four-helix bundle of N-terminal domain (N domain), a Ras-like GTPase domain (G domain), and a binding domain for SRP RNA and the signal peptide (M domain). From crystal structure, the M domain contains a hydrophobic groove as a signal sequence binding pocket, which directly interacts with the hydrophobic stretch in the signal peptide [12, 13].

The preproteins that travel from cytosol to mitochondria also have N-terminal targeting presequence. Preproteins with N-terminal presequences are imported into mitochondria through the TOM (translocases of outer mitochondria membrane) and TIM (translocases of inner mitochondria membrane) complexes. Successive interaction of the presequence with the binding components in TOM and TIM will guide the preprotein to the mitochondrial interior. On the mitochondrial surface, TOM20 is the master receptor for initial recognition of the presequence. Mitochondria matrix-targeting presequences usually contain an amphipathic α–helical conformation where positively charged amino acids predominate on one side and hydrophobic amino acids predominate on the other side of the helix [14]. Deletions or point mutations that change this amphipathic character usually disrupt preprotein import to mitochondria. NMR chemical shift perturbation analysis has shown that Tom20 binds to the hydrophobic surface of the amphipathic α-helix in the presequence [15, 16]. These findings indicate that the amphipathicity of matrix-targeting sequence is

necessary for the mitochondria presequence function [17].

While the amino acid sequences of many chloroplast transit peptides are known, the common features of them are poorly understood [18]. Only a few similarities can be found in chloroplast transit peptide: (i) the second residue is often Ala, (ii) Ser, Leu, Pro, Ala, Gly and Thr are abundant, whereas acidic amino acids are rare. Since the amino acid sequence, length and composition of transit peptide are highly divergent, no consensus in sequence or conformation of transit peptides can be found so far [19, 20].

Previous studies have shown that chloroplast precursor proteins preferred to be imported into younger chloroplasts, and the import capability declined as chloroplasts approach maturity [21]. However, recent data from our lab have shown that chloroplast precursor proteins can be classified into at least three age-selective groups: group I proteins prefer to be imported into younger chloroplasts; import ability of group II proteins is less dependent on chloroplast age; and group III proteins prefer to be imported into older chloroplasts. By swapping the transit peptides of group I and group III proteins, we demonstrated that the age-selective signal is located within the transit peptide of each protein. We have evidence to indicate that Toc159 is the receptor for group I precursors and its homologue Toc132 is the receptor for group II precursors. The receptor for group III precursor is Toc75 in young chloroplasts and is a yet unidentified novel protein in older chloroplasts. We are in the process of identifying this novel receptor. At the same time, it is also important to know what sequence feature in the group III transit peptides that are critical for their recognition of the novel receptor and of older chloroplasts.

In this study, I created many mutant constructs to try to identify the transit peptide sequence motifs important for older chloroplast recognition. I first used the transit peptide of prPsTic40 (precursor of Tic40 from pea *Pisum sativum*), one of the

precursors that prefer to be imported into older chloroplasts as a model. I tried swapping the transit peptide with another homologous transit peptide that prefer younger chloroplasts, or mutating loosely conserved residues in all transit peptides that prefer older chloroplasts. Finally I divided the prPsTic40 transit peptide into eight blocks and generated serial alanine scanning mutants. Combined with several other transit peptide mutation experiments, I found that two consecutive positive charges, arginine and lysine of residues 29 and 30, are necessary and sufficient for older chloroplast recognition. Moreover, in the transit peptide of prAtL11 (precursor of the L11 component of the chloroplast 50S ribosomal subunit from *Arabidopsis thaliana*), another precursor that also prefers to be imported into older chloroplasts, two pairs of two positively charged residues are also responsible for prAtL11 preference for older chloroplasts. Although the exact receptor for prPsTic40 and prAtL11 is still unknown, these data indicate that two positive charges are the critical feature that constitutes the age-selective signal for older chloroplast recognition.

Material and Methods

Plant growth conditions

Pea *(Pisum sativum cv. Little Marvel)* seeds were soaked in running water for 16 hours and grown in vermiculite (size number 4, South Sea Vermiculite and Perlite Co., LTD) under a 12-h light/12-h dark cycle at 20℃ for 16 to 18 days.

PCR-based mutagenesis and construction of plasmids

For alanine scanning mutations, a PCR approach was used to introduce mutations in the transit peptide of prPsTic40 and prAtL11 [22]. For each mutant, one pair of complementary primers was designed. The primers consist of mutated residues in the central region, flanked by wild-type sequences. With the template prPsTic40 and these primers, including the complementary pair, N-terminal and C-terminal primers, the first round of PCR was performed to generate two fragments, the 5' and 3' segments. The second round of PCR was performed with the 5' and 3' segments as templates, and the N-terminal and C-terminal primers. The PCR products were subcloned into pBluescript for the prPsTic40 mutants and pSP72 for the prAtL11 mutants, and then sequenced. For amino acid changes, the QuickChange site-directed mutagenesis kit was used (Agilent Technologies, Inc.). Complementary primers were designed to contain the desired mutations in the middle with 10 to 15 bases of wild-type sequence on both sides. The reaction mixture includes 5 μl of 10X reaction buffer, 50 ng of DNA template, 10 μM of purified primers, 1 μl of dNTP mix, 1 μl of *PfuUltra* HF DNA polymerase (Agilent Technologies, Inc.) and ddH_2O to a final volume of 50 μl. Cycling began at 95℃ for 30 seconds, and then ran at 95℃ for 30 seconds, at 55℃ for 1 minute and at 68℃ for 1 minute/kb of plasmid length for 16 cycles. After the

reaction, 1 μl of the *DpnI* was added to the mixture and incubated at 37℃ for 1 hour. The *DpnI* digested PCR product was then transformed into ECOS101 competent cells and plasmids isolated from the colonies were sequenced.

In vitro transcription and translation

[^{35}S]Met-labeled precursor proteins were obtained from a cell-free wheat germ or reticulocyte lysate extract TNT system (Promega, Madison, WI, USA). For a 100 μl reaction, 50 μl TNT wheat germ or reticulocyte lysate extract, 4 μl TNT reaction buffer, 2 μl RNA polymerase, 2 μl amino acid mix without methionine, 2 μl RNasin, 24 μl [^{35}S]Methionine (Perkin Elmer Life Science, Inc.), 2 μg plasmid DNA encoding precursor proteins and ddH_2O to final volume of 100 μl were mixed and incubated at 30℃ for 2 hours. The translation mixture was then diluted with 100 μl of 2 mM DTT and 50 mM cold methionine in 2X import buffer (1X concentration: 300 mM sorbitol and 50 mM HEPES-KOH, pH8.0). The translation product was examined by SDS-PAGE and stored at -80℃. prHsp93, prOE23, prRBCS, and prAtL11 were translated in wheat germ extracts; prPsTic40 and its mutants were translated in reticulocyte lysates.

Chloroplast isolation

To make a 50% Percoll gradient, 30 ml Percoll (Amersham Biosciences, Uppsala, Sweden) was well mixed with 30 ml 2X pea chloroplasts grinding buffer (600 mM sorbitol, 0.2% BSA, 2 mM $MgCl_2$, 2 mM $MnCl_2$, 4 mM EDTA, 100 mM HEPES-KOH pH 7.3), and then centrifuged at 38000xg at 4℃ for 30 minutes. The first and fourth pair of leaves from 16- to 18-day-old pea seedlings were cut separately from pea seedlings, mixed with grinding buffer and then homogenized with a house blender for 20 to 30 seconds. The homogenate was immediately filtered

through two layers of Miracloth (Calbiochem, Inc.). The filtrate was collected with a beaker on ice and then centrifuged at 3000 xg at 4℃ for 3 minutes. The pellet was resuspended in grinding buffer and the suspension was loaded gently onto the 50% Percoll gradient prepared and the gradient was centrifuged at 7700xg in a swing-bucket rotor at 4℃ for 10 minutes. After the centrifugation, intact chloroplasts, forming a deep green band close to the bottom of the tube, were carefully removed into 30 ml of ice-cold import buffer. The mixture of intact chloroplasts and import buffer was centrifuged at 1500xg at 4℃ for 5 minutes. The pelleted chloroplasts were resuspended in 20 ml import buffer. 10 μl of the resuspension were mixed with 190 μl ddH_20 and 800 μl 100% acetone. The mixture was well mixed and centrifuged at 1500xg at room temperature for 3 minutes. OD_{652} of the supernatant was measured. The chlorophyll concentration was calculated by the following formula: mg chlorophyll/ml=0.0246 x OD_{652} X 0.01. The chloroplast resuspension was spinned again and adjusted to 1 mg chlorophyll/ml with import buffer.

Import assay

For regular import assays, [^{35}S]Met-labeled precursor proteins were incubated with isolated chloroplasts in the presence of 1 mM ATP in import buffer at room temperature for 25 minutes. The import reaction was stopped by diluting with 1 ml ice-cold 40% Percoll (prepared in 1X import buffer) and the intact chloroplasts were pelleted by spinning at 3724 xg for 6 minutes. Chloroplasts were washed with cold import buffer and resuspended in import buffer. Then the samples were mixed with an equal volume of 2X sample buffer (900mM Tris-HCl pH 8.45, 24% glycerol, 8% SDS, 0.0076% Comassie Blue G, 0.01% Phenol Red, 0.001 mM EDTA, 0.1 mM DTT) for SDS-PAGE.

Post-import trypsin treatments

For post-import trypsin treatment, the import assay was stopped by adding excess amount of ice-cold import buffer separated into two halves and spinned down at 3800 rpm for 5 minutes. One half was resuspended in import buffer as the control and the other half was resuspended in import buffer containing 100 μg/ml trypsin (Sigma). Both reactions were incubated in the dark at room temperature for 1 hour. Trypsin digestion was stopped by adding one-tenth volume of 20 mg/ml trypsin inhibitor (Sigma). The reaction mixture was kept in the dark on ice for ten minutes. Intact chloroplasts were reisolated and processed for SDS-PAGE as described above.

Gel electrophoresis and immunoblotting

Electrophoresis was carried out using the NuPAGE® Novex® 4-12% Bis-Tris Gels (Invitrogen) with MES running buffer (50 mM MES, 50 mM Tris base, 0.1% SDS, and 1 mM EDTA). Gels were run at a constant voltage (140 V) for approximately 1 hr. After electrophoresis, gels were soaked in a Comassie blue staining solution (0.1% Coomassie Brilliant Blue R-250, 50% methanol and 10% glacial acetic acid) for 20 minutes with gentle shaking and then destained in a destaining solution (40% methanol and 10% glacial acetic acid). The destained solution was changed several times until the background of the gels was clear and the gels were then moved to water to wash away all the destain solution. Gels were then dried for 40 minutes on a gel dryer. Dried gels were exposed to Fuji imaging plates for quantification of [^{35}S]Met-labeled protein bands using the Fuji FLA5000 phosphor imager (Fuji Photo Film) or exposed to Biomax MS film with TranScreen (Kodak). Immunoblots were performed by electroblotting of proteins onto polyvinylidene difluoride (PVDF) membranes, probing membranes using specific primary antibodies and an alkaline phosphatase-conjugated secondary antibody and visualized by the NBT-BCIP colorimetric system.

Results

Determine the usable sites for crosslinkers

We intend to use p-benzoyl-L-phenylalanine (Bpa) mediated photocrosslinking to identify the novel receptor specific for group III precursor proteins in older chloroplasts. This method uses the non-natural amino acid Bpa. Bpa residue is incorporated into the nascent polypeptide chain by the mutated tyrosine tRNA synthase that recognizes the amber TAG stop codon [23]. However, we need to first verify that the substitution of Bpa at a particular residue would not significantly alter the age preference of the prPsTic40 transit peptide. I thus created site-directed mutagenesis mutants, in which every tenth residue in the transit peptide was substituted to tyrosine (Tyr) and then checked the age preference of these mutants in import assays (Fig. 1A). Our previous results have shown that the precursor to the small subunit of RuBP carboxylase (prRBCS) is a group I precursor that prefers to be imported into younger chloroplasts, and prHsp93 is a group II precursor that shows similar import efficiency into chloroplasts isolated from all four leaves. Thus, in all the import experiments of this thesis, prHsp93 was co-imported with each precursor as an internal control. The import results were normalized to the amount of mature Hsp93 imported. As shown in Fig. 1B and C, compared with wild-type prPsTic40, most of the mutants had a similar preference for older chloroplasts, except maybe the 30Y mutant, which had a slightly lower preference for older chloroplasts. Therefore, we should avoid adding crosslinker to this position for future experiments.

The crucial region for older chloroplast recognition is at the N-terminal portion of prPsTic40 transit peptide

From our previous results, prPsTic40 prefers to be imported into older chloroplasts while prPpTic40 from the moss *Physcomitrella patens* prefers to be imported into younger chloroplasts. To identify the region in the prPsTic40 transit peptide responsible for older chloroplast recognition, the transit peptide was first divided into two parts and two fusion constructs were produced by Dr. Yi-Shan Teng. From the predicted secondary structure, prPsTic40 transit peptide contains an α-helix structure at residues 45 to 54 but prPpTic40 does not. To preserve this possible α-helix structure in the prPsTic40 transit peptide, in the first fusion construct, prPsTic40(1-54)-PpTic40, the N-terminal 54 residues of prPpTic40 transit peptide were replaced by the N-terminal 54 residues of prPsTic40 transit peptide (Fig. 2A). In the second construct, prPpTic40(1-55)-PsTic40, the N-terminal 55 residues of prPsTic40 transit peptide were replaced by the N-terminal 55 residues of prPpTic40 transit peptide (Fig. 2A). I then performed import assays using both the wild types and the fusion constructs (Fig. 2B). Similar to previous data, prPpTic40 and prPsTic40 preferred younger and older chloroplasts, respectively. Furthermore, prPpTic40(1-55)-PsTic40 was preferentially imported into younger chloroplasts, like prPpTic40, and prPsTic40(1-54)-PpTic40 was preferentially imported into older chloroplasts, just like prPsTic40 (Fig. 2B and C). These results indicated that the N-terminal 55 residues in the prPsTic40 transit peptide are important for older chloroplast recognition.

The older chloroplast preference was not changed by single residue substitutions in the transit peptide of prPsTic40

Since transit peptides are sufficient to determine the age selectivity, the sequence differences between the transit peptide of prPsTic40 and prPpTic40 were examined. The protein sequences in Fig. 3A are from group III precursors; even all of them

prefer older chloroplasts, their primary sequences and predicted secondary structure are highly divergent. The alignment of Fig. 3B is the same as Fig. 3A except the sequence of prPpTic40, a group I precursor, was added. When prPpTic40 was added into the alignment, two residues of conservation in group III transit peptide disappeared: a place where all the group III transit peptide have an amino acid with a polar side chain (pointed by red arrow, Fig. 3A and B) and the other where all have an amino acid with a hydrophobic side chain (pointed by blue arrow, Fig. 3A and B). To test whether the two residues are critical for older chloroplast recognition, the two residues, serine of the thirty-first residue (S31) and valine of the sixty-second residue (V62), were mutated. They were mutated to corresponding residues of the prPpTic40 transit peptide and created mutants S31G and V62S. In addition, S31 was also mutated to Pro (S31P) and V62 was mutated to Glu (V62E, Fig. 3C). Ser has a flexible side chain. Mutating it to Pro should change the original protein structure. Val has a nonpolar and hydrophobic side chain, so mutating it to Glu created a hydrophobicity change. Import assays showed that these four mutants had import patterns that were nearly identical to the wild type, indicating that they were not crucial for the age selectivity.

The crucial region of prPsTic40 is at residues 28 to 36

To more systematically identify the regions in the prPsTic40 transit peptide responsible for older chloroplast recognition, a series of alanine (Ala) substitution mutants were generated. The prPsTic40 transit peptide was divided into eight blocks of nine amino acids, and each block was replaced with nine Ala residues (Fig. 4A). Ala is the substitution residue of choice because it eliminates the side chain beyond the β carbon and neither alters the main-chain conformation nor does it impose extreme electrostatic or steric effects [24]. These Ala-substitution mutants were

imported into young and old chloroplasts. All mutants, with the exception of A(10-18) and A(28-36), still showed a preference for older chloroplasts (Fig. 4B and C), indicating that these mutant transit peptides were able to be recognized by the novel group III receptor. The A(28-36) mutation did not affect the import into young chloroplasts but severely reduced the import into older chloroplasts, indicating that residues 28 to 36 are important for older chloroplast recognition.

No mature protein of the A(10-18) mutant was observed after its import (Fig. 4B). It is possible that the A(10-18) mutant precursor could only bind but could not be translocated into chloroplasts. Or the A(10-18) mutant precursor might have been translocated into chloroplasts but the mutation had abolished transit peptide cleavage by SPP. Because trypsin can only penetrate the outer membrane but not the inner membrane, it can be used to distinguish whether A(10-18) mutant was outside the inner membrane or in the stroma. I thus treated chloroplasts after import of A(10-18) with trypsin. As shown in Figure 5A, the A(10-18) mutant precursor was nearly completely degraded by the trypsin post-treatment while wild-type prPsTic40 was resistant to trypsin digestion, indicating that A(10-18) mutant was localized outside the inner membrane. Western-blot analyses of Toc75, an integral outer-membrane protein, and Tic110, an integral inner-membrane protein, were also performed to confirm that the trypsin post-treatment had been performed properly (Fig. 5B). These results suggest that residues of 10 to 18 of prPsTic40 transit peptide may be recognized by Toc75, the channel for precursors crossing the outer membrane.

The crucial region of prPsTic40 is at residues 28 to 31

To further narrow down the region responsible for the age selectivity, residues 28 to 36 were further divided into two regions: the former four residues and the latter five residues, and each region was substituted with corresponding numbers of Ala

(Fig. 6A, A(28-31) and A(32-36)). I have also noticed an interesting feature around this region: there are four phenylalanine (Phe) residues in residues 26 to 36. Because Phe is not an abundant residue in transit peptides of chloroplast precursor proteins, I suspected that Phe may also play an important role for the age selectivity. Therefore, Ala substitutions were extended to residues 26 and 36 and three more mutants, A(26-31), A(32-37) and A(31-35), were also generated for assays (Fig.6A). Among these five mutants, A(28-31) and A(26-31) no longer preferred older chloroplasts while the age preference of the other three mutants was the same as the wild type (Fig. 6B and C). Since mutation in residues 28 to 31 was sufficient to abolish older chloroplast preference, I concluded that, residues 28-31, Gly-Arg-Lys-Ser, were critical for interacting with the novel receptor for group III precursors in older chloroplasts.

Arginine 29 and lysine 30 of prPsTic40 is important for older chloroplast recognition

Gly-Arg-Lys-Ser of residues 28 to 31 were further divided into three regions, Gly, Arg-Lys, and Ser. These residues were substituted by not only Ala but also the amino acids with the most different characteristics. Gly with the smallest side group was substituted by Pro and His. Gly is the most flexible residue while Pro has stronger stereochemical constraints and His has a heterocyclic aromatic amine side chain. Pro is both disruptive to regular secondary structure and good at forming turns in polypeptide chains. Arg and Lys, which are positively charged, were substituted by Glu due to its negative charge. Because of its hydroxyl and uncharged side chain, Ser was also substituted by Glu (Fig. 7A). As shown in Fig 7B and C, with the exception of the Arg-Lys mutants, all other mutants had an age preference pattern that was nearly identical to the wild type. The RK2930EE mutant had almost lost the ability to

be imported into older chloroplasts, very similar to the A(28-36) mutant (Fig. 4C). These results indicated that the two residues, arginine 29 (R29) and lysine 30 (K30), are essential for older chloroplast recognition.

Two consecutive positively charged residues in the prPsTic40 transit peptide are necessary and sufficient for older chloroplast recognition

To test whether both R29 and K30 are crucial for older chloroplast recognition, they were each substituted with Glu (Fig. 8A). As shown in Fig 8B and C, preference for older chloroplasts was reduced, though not abolished, in the R29E and K30E mutants. These results indicated that both R29 and K30 are required to confer full older chloroplast preference.

To further examine if R29 and K30 were sufficient to confer older chloroplast preference, they were re-added to the A(28-31), A(26-31) and A(28-36) mutants, which have lost the preference for older chloroplasts (Fig. 8A). As shown in Fig. 8B and C, the three new mutants, A(28-31)+RK, A(26-31)+RK and A(28-36)+RK, regained their preference for older chloroplasts. These results indicated that R29 and K30 were necessary and sufficient for older chloroplast preference.

Two positively charged residues in the prAtL11 transit peptide are responsible for older chloroplast recognition

The results obtained from the import experiments with the prPsTic40 mutants indicate that two consecutive positively charged residues were crucial for older chloroplast recognition. To confirm this observation, mutants of prAtL11, another precursor that also prefers to be imported into older chloroplasts, were generated. The transit peptide of prAtL11 has two consecutive positively charged residues at residues 44 and 45. These two residues were therefore mutated to Glu or Ala, creating mutants

prAtL11(KK4445EE) and prAtL11(KK4445AA). In addition, there is a predicted β-strand structure located in residues 31 to 41 in the transit peptide of prPsTic40, just after the crucial R29 and K30. There is a predicted β-strand structure located in residues 36 to 41 in the transit peptide of prAtL11. Preceding this β-strand structure, there are also two positively charged residues, K30 and K34, although these two residues are separated by three other amino acids. I decided to also mutate them, together or K34 alone, into Ala or Glu, and created mutants prAtL11(K34A), prAtL11(K34E), prAtL11(KK3034AA), and prAtL11(KK3034EE) (Fig. 9A). As shown in Fig. 9B and C, the K34A and K34E mutations had little effect in the age preference of prAtL11. Mutating either pair of lysines to Ala (KK3034AA and KK4445AA) reduced the preference for older chloroplasts and mutating them to Glu (KK3034EE and KK4445EE) almost abolished import into older chloroplasts. The import efficiencies of mutants into younger or older chloroplasts were further compared to the wild-type prAtL11 (Fig. 9D). Compared to prAtL11 import efficiency in older chloroplasts, the import efficiency of all mutants into older chloroplasts was decreased. The import efficiency of mutants into younger chloroplasts, with the exception of KK4445EE, was also decreased (Fig. 9D). These data indicated that there are two pairs of positively charged residues in the prAtL11 that are important for older chloroplast import. The mutation in KK3034EE affected import both in younger and older chloroplasts. The mutation in KK4445EE only affected import into older chloroplasts but not import into younger chloroplasts.

Discussion

The aim of this thesis is to identify the sequence motifs in the transit peptides of prPsTi40 and prAtL11 critical for age selectivity. The transit peptide of prPsTic40 was examined first. Among the nine-amino acid block mutants, Ala substitution in residues 28 to 36 was the most detrimental for older chloroplast recognition (Fig. 4). Import assays using additional Ala substitution mutants further showed that the crucial region is at the region of residues 28 to 31(Fig. 6). My data demonstrated that residues R29 and K30 of the prPsTic40 transit peptide are necessary and sufficient for the age selectivity (Fig. 7 and 8). Therefore, we hypothesized that two consecutive positively charged residues were responsible for older chloroplast recognition. In support of our hypothesis, two pairs of positively charged residues in the prAtL11 transit peptide were mutated and the result showed that mutations in both pairs affected import into older chloroplasts (Fig. 9).

If the novel receptor specific for group III precursors has a negatively charged pocket to recognize the two positively charged residues in the transit peptide (Fig. 10), it is hard to explain why prPsTic40 only has one pair but prAtL11 has two. From the import efficiency data of young chloroplasts, an interesting difference existed between the ptAtL11 mutants KK3034EE and KK4445EE. Although their import efficiencies into older chloroplasts decreased dramatically, the import efficiency of KK4445EE into younger chloroplasts was almost identical to that of wild type but KK3034EE had decreased import efficiency even into younger chloroplasts (Fig. 9D). Based on the data that Toc75 is the channel across the outer membrane for most precursor translocation [25, 26], I propose that K44 and K45 in the prAtL11 transit peptide may be recognized by the novel receptor and mutations in K30 and K34 may have affected interaction with Toc75. In other words, when K44 and 45 were mutated, prAtL11 was

not recognized by the novel receptor but could still normally interact with Toc75; thus, import efficiency into younger chloroplasts was not affected. On the other hand, when K30 and K34 were mutated, prAtL11 was not recognized by Toc75, so that both import efficiency into younger and older chloroplast noticeably declined.

Through serial Ala substitutions, several sequence motifs in the prRBCS transit peptide have been identified to be important for chloroplast targeting [27]. Among the Ala substitution blocks, the region of residues 32 to 41 is most critical for targeting to chloroplasts, and the block contains two important motifs, Phe-Pro (FP) and Arg-Lys (RK). Import efficiency was fully restored by adding FP and partially restored by adding RK back to the mutant in which residues 32 to 41 were substituted with alanines. However, it is not clear why both FP and RK could restore the import efficiency. The position of this block, in the middle of prRBCS transit peptide, is similar to our prPsTic40 result. Since only R29 and K30 contribute to prPsTic40 age selectivity, it seems that the necessary sequence motifs in prRBCS transit peptide are more complicated and have more redundancy than that of prPsTic40.

The results in this thesis indicate that positively charged residues likely serve as binding sites in prPsTic40 and prAtL11 for the novel receptor on older chloroplasts. As shown in the transit peptide sequence alignment in Fig. 3A, even though the positively charged residues are located at different positions, AtTic40 has two Arg at residues 34 and 35, OsTic40 has two Arg at residues 41 and 42, and SmTic40 also has two Arg at residues 35 and 36, suggesting that two consecutive positively charged residues may be crucial for older chloroplast recognition. However, prRBCS, a precursor that prefers to be imported into younger chloroplasts, also has Arg and Lys at residues 37 and 38; prPpTic40, another precursor that also prefers to be imported into younger chloroplasts, has three pairs of positively charged residues at 45 and 48, 51 and 53, 64 and 65, respectively. It is possible that the presence of two consecutive

positively charged residues is only one of the necessary features. Other yet unknown structure motifs in the prPsTic40 and prAtL11 transit peptide are required to work together with the consecutive positively charged residues.

Because we have so far only identified two proteins that prefer to be imported into chloroplasts, it is difficult to know what kind, in terms of physiological functions, of proteins need to be preferentially imported into older chloroplasts. The next step will be to identify more group III precursors and to identify the novel receptor that recognizes the group III precursors. Further studies will also be necessary to define the exact roles of those critical residues on how and where they recognize, bind and interact with the receptor.

References

1. Howe CJ, Barbrook AC, Nisbet RE, Lockhart PJ, Larkum AW: **The origin of plastids**. *Philos Trans R Soc Lond B Biol Sci* 2008, **363**(1504):2675-2685.
2. Leister D: **Chloroplast research in the genomic age**. *Trends Genet* 2003, **19**(1):47-56.
3. van Wijk KJ, Baginsky S: **Plastid proteomics in higher plants: current state and future goals**. *Plant Physiol* 2011, **155**(4):1578-1588.
4. Keegstra K: **Transport and routing of proteins into chloroplasts**. *Cell* 1989, **56**(2):247-253.
5. Bauer J, Chen K, Hiltbunner A, Wehrli E, Eugster M, Schnell D, Kessler F: **The major protein import receptor of plastids is essential for chloroplast biogenesis**. *Nature* 2000, **403**(6766):203-207.
6. Kessler F, Schnell DJ: **The function and diversity of plastid protein import pathways: a multilane GTPase highway into plastids**. *Traffic* 2006, **7**(3):248-257.
7. Li H-m, Chiu C-C: **Protein Transport into Chloroplasts**. *Annual Review of Plant Biology* 2010, **61**(1):157-180.
8. Schnell DJ, Blobel G, Keegstra K, Kessler F, Ko K, Soll J: **A consensus nomenclature for the protein-import components of the chloroplast envelope**. *Trends Cell Biol* 1997, **7**(8):303-304.
9. Keegstra K, Froehlich JE: **Protein import into chloroplasts**. *Curr Opin Plant Biol* 1999, **2**(6):471-476.
10. Wickner W, Schekman R: **Protein translocation across biological membranes**. *Science* 2005, **310**(5753):1452-1456.
11. Walter P, Johnson AE: **Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane**. *Annu Rev Cell Biol* 1994, **10**:87-119.
12. Rapoport TA: **Transport of proteins across the endoplasmic reticulum membrane**. *Science* 1992, **258**(5084):931-936.
13. Montoya G, Kaat K, Moll R, Schafer G, Sinning I: **The crystal structure of the conserved GTPase of SRP54 from the archaeon Acidianus ambivalens and its comparison with related structures suggests a model for the SRP-SRP receptor complex**. *Structure* 2000, **8**(5):515-525.
14. Roise D, Horvath SJ, Tomich JM, Richards JH, Schatz G: **A chemically synthesized pre-sequence of an imported mitochondrial protein can form an amphiphilic helix and perturb natural and artificial phospholipid bilayers**. *EMBO J* 1986, **5**(6):1327-1334.
15. Igura M, Ose T, Obita T, Sato C, Maenaka K, Endo T, Kohda D: **Crystallization**

and preliminary X-ray analysis of mitochondrial presequence receptor Tom20 in complexes with a presequence from aldehyde dehydrogenase. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 2005, **61**(Pt 5):514-517.

16. Muto T, Obita T, Abe Y, Shodai T, Endo T, Kohda D: **NMR identification of the Tom20 binding segment in mitochondrial presequences**. *J Mol Biol* 2001, **306**(2):137-143.
17. Roise D, Theiler F, Horvath SJ, Tomich JM, Richards JH, Allison DS, Schatz G: **Amphiphilicity is essential for mitochondrial presequence function**. *EMBO J* 1988, **7**(3):649-653.
18. Bruce BD: **The paradox of plastid transit peptides: conservation of function despite divergence in primary structure**. *Biochim Biophys Acta* 2001, **1541**(1-2):2-21.
19. von Heijne G, Steppuhn J, Herrmann RG: **Domain structure of mitochondrial and chloroplast targeting peptides**. *Eur J Biochem* 1989, **180**(3):535-545.
20. Zybailov B, Rutschow H, Friso G, Rudella A, Emanuelsson O, Sun Q, van Wijk KJ: **Sorting signals, N-terminal modifications and abundance of the chloroplast proteome**. *PLoS ONE* 2008, **3**(4):e1994.
21. Dahlin C, Cline K: **Developmental Regulation of the Plastid Protein Import Apparatus**. *Plant Cell* 1991, **3**(10):1131-1140.
22. Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR: **Site-directed mutagenesis by overlap extension using the polymerase chain reaction**. *Gene* 1989, **77**(1):51-59.
23. Ryu Y, Schultz PG: **Efficient incorporation of unnatural amino acids into proteins in Escherichia coli**. *Nature Methods* 2006, **3**(4):263-265.
24. Lefevre F, Remy MH, Masson JM: **Alanine-stretch scanning mutagenesis: a simple and efficient method to probe protein structure and function**. *Nucleic Acids Res* 1997, **25**(2):447-448.
25. Hinnah SC: **The chloroplast protein import channel Toc75: pore properties and interaction with transit peptides**. *Biophys J* 2002, **83**(2):899-911.
26. Schleiff E, Soll J, Kuchler M, Kuhlbrandt W, Harrer R: **Characterization of the translocon of the outer envelope of chloroplasts**. *J Cell Biol* 2003, **160**(4):541-551.
27. Lee DW: **Functional Characterization of Sequence Motifs in the Transit Peptide of Arabidopsis Small Subunit of Rubisco**. *Plant Physiology* 2006, **140**(2):466-483.