Differential usage of two in-frame translational start codons regulates subcellular localization of *Arabidopsis thaliana* THI1

Sabrina M. Chabregas[1], Douglas D. Luche[2], Marie-Anne Van Sluys[3], Carlos F. M. Menck[2] and Marcio C. Silva-Filho[1,*]

[1]Departamento de Genética, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Caixa Postal 83, 13400-970, Piracicaba, SP, Brazil
[2]Departamento de Microbiologia, Instituto de Ciências Biomédicas, Universidade de São Paulo, Av. Prof. Lineu Prestes 1374, 05508-900, São Paulo, SP, Brazil
[3]Departamento de Botânica, Instituto de Biociências, Universidade de São Paulo, Caixa Postal 11461, 05422-970, São Paulo, SP, Brazil
*Author for correspondence (e-mail: mdcsilva@esalq.usp.br)

Accepted 15 October 2002
Journal of Cell Science 116, 285-291 © 2003 The Company of Biologists Ltd
doi:10.1242/jcs.00228

Summary

***Arabidopsis thaliana* THI1 is encoded by a single nuclear gene and directed simultaneously to mitochondria and chloroplasts from a single major transcript. In vitro transcription/translation experiments revealed the presence of two translational products by the differential usage of two in-frame translational start codons. The coupling site-specific mutations on the THI1 encoding sequence with green fluorescent protein (GFP) gene fusions showed that translation initiation at the first AUG directs translocation of THI1 to chloroplasts. However, when translation starts from the second AUG, THI1 is addressed to mitochondria. Analysis of the translation efficiency of *thi1* mRNA revealed that the best context for translation initiation is to use the first AUG. In addition, a suboptimal context in the vicinity of the second AUG initiation codon, next to a stable stem-and-loop structure that is likely to slow translation, has been noted. The fact that translation preferentially occurs in the first AUG of this protein suggests a high requirement for TH1 in chloroplasts. Although the frequency of upstream AUG translation is higher, according to the first AUG rule, initiation at the second AUG deviates significantly from Kozak's consensus. It suggests leaky ribosomal scanning, reinitiation or the internal entry of ribosomes to assure mitochondrial protein import.**

Key words: Dual targeting, THI1, Mitochondria, Chloroplasts, Import, Translational regulation

Introduction

Multicompartmentalized proteins encoded by a single gene originate by different mechanisms operating at both transcriptional and post-transcriptional levels (for a review, see Danpure, 1995; Peeters and Small, 2001). An increasing number of examples of dual targeted proteins has been reported in fungi and higher eukaryotes. The two most common results of variable expression of N-terminal targeting sequences are alternative forms of transcription and translation initiation (Gebhardt et al., 1998), although the molecular mechanisms underlying these processes have not yet been fully elucidated (Sass et al., 2001).

The occurrence of mitochondria and chloroplasts in plant cells requires a higher organellar protein import specificity than in non-plant sources (Glaser et al., 1998). Thus, protein targeting to both endosymbiotic organelles follows distinct pathways and involves different targeting sequences and import machineries; this underlines the complexity of the relationship between these cell compartments (Duchêne et al., 2001; Gabriel et al., 2001). Although some mistargeting has been reported so far (Franzén et al., 1990; Hurt et al., 1986; Silva-Filho, 1999; Whelan et al., 1990), this is because of heterologous expression systems or unusual targeting sequences. In homologous systems mistargeting seems not to occur in vivo (Silva-Filho et al., 1997); therefore, dual targeting to mitochondria and chloroplasts is apparently not related to mistargeting but to cell requirement (Soll and Tien, 1998). An interesting report has shed light on the protein import specificity of both organelles. It has been suggested that a mitochondrial preprotein receptor, called Tom22, is required for cohabitation of mitochondria and plastids in the same cell, in such a way as to prevent mistargeting of chloroplast proteins (Macasev et al., 2000).

The manner in which the proteins targeted to both mitochondria and chloroplasts are recognized by both import machineries is still poorly understood. It has been proposed that most of these proteins carry ambiguous targeting signals (Hedtke et al., 2000; Small et al., 1998). However, recent reports showed that targeting of spinach protoporphyrinogen oxidase II (Watanabe et al., 2001) and of a phage-type RNA polymerase (Kobayashi et al., 2001) to both mitochondria and chloroplasts is mediated by the alternative use of two in-frame initiation codons.

The single copy *thi1* gene encodes the bivalent protein THI1 protein, which is targeted both to chloroplasts and mitochondria (Chabregas et al., 2001). This protein plays a role

in the biosynthesis of thiamine (vitamin B1) and may be involved in protection against organellar DNA damage (Machado et al., 1996; Machado et al., 1997). The mature THI1 protein is synthesized with a typical N-terminal chloroplastic transit peptide, which is in agreement with the plastid location for the pathway (Belanger et al., 1995). We have previously shown the presence of a sequence adjacent to the chloroplast-targeting signal that is able to fold into an amphiphilic α-helix. This structure was shown to be involved in THI1 mitochondrial import as well as in targeting of a reporter protein into the organelle (Chabregas et al., 2001). Furthermore, THI1 is synthesized from a single nuclear transcript, which suggests that a post-transcriptional mechanism is responsible for the final localization of the protein in *Arabidopsis*.

In the present study, we examine in greater detail the mechanism responsible for dual targeting of THI1 to mitochondria and chloroplasts. Interestingly, two translational products appeared from *thi1* mRNA by the use of two in-frame AUG codons. Translation initiation efficiency was significantly higher around the first AUG, which in fact presents the optimum context. Surprisingly, translation initiation still occurred at the second AUG codon notwithstanding a poor context and a strong stem-and-loop structure. Alteration of the translation initiation context around the two AUG codons by site-directed mutagenesis significantly affected the translation initiation, especially for the first AUG. In addition, we have prepared a set of truncated constructs by the fusion of THI1 transit signals to the green fluorescent protein. The results show that the distribution of THI1 in *A. thaliana* is determined by a differential usage of the translational initiation codons. Therefore, the evidence presented indicates that the less favorable context for translation initiation at the second AUG is used to direct translocation of THI1 to mitochondria.

Materials and Methods

In vitro transcription/translation of thi1-gus constructs

Standard procedures were used for recombinant DNA work (Sambrook and Russel, 2001). Six constructions corresponding to the entire THI1 transit peptide followed by 64 amino acids from mature THI1, including the mitochondrial presequence, were obtained from *thi1* cDNA cloned into the *Eco*RI restriction site of plasmid KS– (Machado et al., 1996). For amplification of the THI64 fragment, the following set of primers was used: 5′-CCCGGTACCCAAAATGGCTGC-3′ and 5′-CCCAAGCTTCCAGGACTAACAGATTG-3′. The first primer introduced a *Kpn*I site in front of the *thi1* coding sequence, and the second primer introduced a *Hind*III site in a region corresponding to the 64th amino acid of the mature THI1, producing the THI64. In order to create the construct THI64Δ1, the first ATG initiation codon was mutated to ATC with the oligonucleotide 5′-CCCGGTACCCAAAATCGCTGC-3′ (mutation underlined). To create the construction THI64Δ2, in which the second ATG (Met-70) of THI1 is converted into an ATC codon (Ile), two PCR products were obtained. The first 212 *thi1* bp fragment (named thia) was obtained from the upstream primer: 5′-CCCGGTACCCAAAATGGCTGC-3′ carrying a *Kpn*I site and the downstream primer 5′-CCCC-TCGAGACGATCGATTCCTTG-3′, which presents a *Xho*I site. The second 155 bp fragment (named thibΔ) carrying the (Met70Ile) mutation (underlined) was obtained from the primer 5′-CCCC-TCGAGGTCACGTGAGATCACGAG-3′ carrying a *Xho*I site and the primer 5′-CCCAAGCTTCCAGGACTAACAGATTG-3′, which carries a *Hin*dIII site. The two DNA fragments were cut with *Xho*I and joined back with T_4 DNA ligase, producing the construct THI64Δ2.

In order to modify the translation initiation context around the two in-frame AUG codons, three constructions were obtained by PCR-directed mutagenesis. In the first construct, named THI64-1c, the AAA*ATG*GCT context in the vicinity of the first ATG (italics) was mutated to CAA*ATG*CCT (mutations underlined). The upstream primer carrying a *Kpn*I site was 5′-CCCGGTACCCCAAATGCCT-GC-3′ and the downstream primer which carries a *Hin*dIII site was 5′-CCCAAGCTTCCAGGACTAACAGATTG-3′. To obtain the second construct, named THI64-2c, in which the context around the second in-frame ATG (italics) GAG*ATG*ACG was modified to ACC*ATG*GCG (mutations underlined), two PCR fragments were brought together. The first 212 bp fragment obtained (thia) has already been described. The second fragment (called thib), corresponding to 155 bp, was obtained from the primer 5′-CCCCTCGAGGTCACGTACCATG-GCGAG-3′, carrying a *Xho*I site, and the primer 5′-CCC-AAGCTTCCAGGACTAACAGATTG-3′, which carries a *Hin*dIII site. These fragments were cut with *Xho*I and ligated with T4 DNA ligase. Finally, in the third construct, called THI64-12c, both initiation codons were altered (same modifications as described before). In this case, a fragment similar to thia (named thiaΔ) obtained with the upstream primer 5′-CCCGGTACCCCAAATGCCTGC-3′ and downstream primer 5′-CCCCTCGAGACGATCGATTCCTTG-3′ and thib were joined.

All constructs were engineered to contain *Kpn*I and *Hin*dIII sites at the 5′ and 3′ ends, respectively, in order to clone them in-frame into the respective sites of the plasmid SK+ GUS (Silva-Fillho et al., 1996). The resulting plasmids were as follows: SK+ THI64-GUS, SK+ THI64Δ1-GUS, SK+ THI64Δ2-GUS, SK+ THI64-1c-GUS, SK+ THI64-2c-GUS and SK+ THI64-12c-GUS.

In-vitro-coupled transcription and translation were performed using the T7 TNT Quick-Coupled Transcription/Translation System (PROMEGA), according to the manufacturer's instructions, in the presence of [^{35}S] methionine (Amersham Pharmacia). One tenth of the translated proteins was run on 10% SDS-PAGE and dried before exposure. Signals were quantified using the GST-700 Imaging Densitometer and the software Molecular Analyst (BioRad).

Transport experiments using GFP fusion proteins

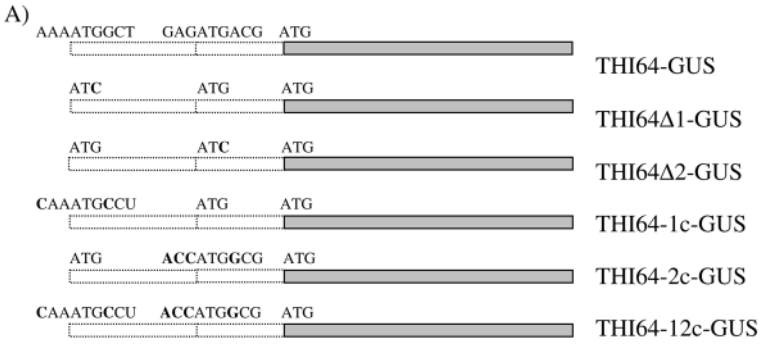
The sequence corresponding to the N-terminus of THI1 (nucleotides 1-315), which includes the entire length of the chloroplastic transit peptide followed by 50 amino acids of the mature THI1 (including the putative mitochondrial presequence), was amplified by PCR using the *A. thaliana* cDNA as a template. A similar approach to the *thi1gus* constructions (see below) was used to generate the constructs named THI50 (two ATGs in-frame), THI50Δ1 (first ATG was mutated to ATC) and THI50Δ2 (second in-frame ATG was mutated to ATC), similar to THI64, THI64Δ1 and THI64Δ2, respectively. Furthermore, an additional construct starting from the second in-frame ATG, named THImet2, was also generated by PCR. The plasmid SK+ THI64Δ2-GUS was used as a template to obtain the THI50Δ2 fragment. For all PCR reactions the downstream primer was 5′-CCCACTAGTA-GGGTTCTTACTGATCTC-3′. The THI50 and THI50Δ2 upstream primers were 5′-CCCAGATCTCAAAATGGCTGC-3′. To obtain the THI50Δ1 and THImet2 fragments the upstream primers were, respectively, 5′-CCCAGATCTCAAAATCGCTGC-3′ and 5′-CCCA-GATCTGAGATGACGAGAAGGTAC-3′. All DNA fragments were engineered to contain *Bgl*II and *Spe*I restriction sites at the 5′ and 3′ end, respectively. These fragments were cloned into the GFP expression vector pCAMBIA1302 (Roberts et al., 1997), previously digested by *Bgl*II and *Spe*I. The resulting vectors were designated to as pCambia-THI50-GFP, pCambia-THI50Δ1-GFP, pCambia-THI50Δ2-GFP and pCambia-THImet2-GFP.

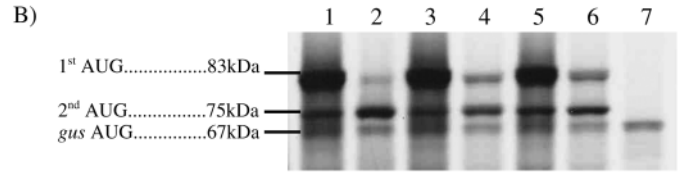
These constructions were introduced into tobacco mesophyll protoplasts, prepared as follows: Green leaves of *Nicotiana tabacum*

(var. SRI), grown in vitro (~250 mg), were cut into small strips using a new razor blade and incubated in 5 ml of an enzyme solution [0.2% Macerozyme R-10 (Yakult Honsha Co., Ltd., Tokyo, Japan), 1.0% Cellulase Onozuka R-10 (Merck), 0.5% Driselase (Sigma) in CPW8 (Frearson et al., 1973)] at room temperature for 16 hours under gentle agitation (~30-40 rpm). After incubation, the protoplast suspension was filtered through a 64 μm mesh, the protoplasts being collected by centrifugation at 46 ***g*** for 5 minutes. The pelleted protoplasts were resuspended in 5 to 10 ml of CPW8 solution and centrifuged. After centrifugation, the protoplasts were resuspended in MKCl (500 mM mannitol, 5 mM KCl and 200 μM MOPS) at a density of 2×10^6 protoplasts/ml. Protoplasts were electroporated with 10-50 μg of the respective plasmid DNAs and incubated in the dark for 24-48 hours before analysis. Transient expression of the GFP constructs in eletroporated protoplasts was analyzed using an epifluorescence microscope Axioplan 2 (Zeiss). For GFP fluorescence, excitation was at 450-490 nm and emission at 520 nm. Chloroplast autofluorescence was detected between 664 and 696 nm with an excitation at 488 nm.

mRNA structure analysis

The secondary structure prediction for *thi1* mRNA was performed by informatic tools directly available at the http://mfold2.wustl.edu/~mfold/rna/for1-2.3.cgi and http://www.ibc.wustl.edu/~zuker/rna/ sites.





C)

Lane	1	2	3	4	5	6
1st AUG	4.75	1.02	3.88	1.54	4.50	1.58
2nd AUG	1.49	2.87	1.29	1.93	1.75	2.31
gus	1.00	1.00	1.00	1.00	1.00	1.00
ratio	**3.19**	**0.35**	**3.00**	**0.80**	**2.57**	**0.68**

Fig. 1. Schematic representation of the chimeric gene constructs used for the in vitro transcription/translation experiments and functional in vitro analysis of the translation initiation context in the vicinity of the two potential in-frame AUG start codons in the *thi1* mRNA. (A) The THI64-GUS construct retains both 5′ *thi1* in-frame ATGs intact codons. THI64Δ1-GUS contains a mutated first codon and an intact second codon. THI64Δ2-GUS presents an intact first ATG codon and a mutated second codon. THI64-1c-GUS contains a poor context introduced around the first ATG start codon. THI64-2c-GUS presents an optimum context inserted in the vicinity of the second in-frame ATG. In THI64-12c-GUS both initiation codons were altered (same modifications as described before). (B) Translational products from the first and second in-frame AUGs, 83 kDa and 75 kDa, respectively, of the chimeric gene constructions presented above (lanes 1 to 6) and SK+ GUS (67 kDa) used as a control. (C) Quantification of the translational products of the chimeric gene constructs. The GUS translation initiation efficiency was assumed to be similar in all situations. The intensity of the translational products is given as the ratio of the translation initiation in each of the two in-frame AUG start codons to that observed at the GUS initiation codon.

Results

THI1 presents two sites of translation initiation

thi1 presents two in-frame AUG codons both with the potential to be translation initiation sites. The functionality of the two *thi1* in-frame AUG codons was investigated by analysis of the protein products from constructions of the THI1 N-terminal portion fused to the *GUS* gene, generated by in-vitro-coupled transcription/translation using a rabbit reticulocyte lysate, obtained by the use of the TNT T7 Quick Coupled Transcription/Translation System. Thus, we prepared gene constructs assembling the two putative *thi1* in-frame initiation codons and the *GUS* reporter gene bearing its own AUG codon for translation initiation (Fig. 1A). These chimeric genes were placed in an *Escherichia coli* vector, under the control of a T7 RNA polymerase promoter. The first construct, THI64-GUS (Fig. 1A), retained the entire chloroplastic transit peptide followed by 64 amino acids of mature THI1 (including the putative mitochondrial presequence). When it was used as a template, two translational products with apparent molecular masses of 83 and 75 kDa were detected (Fig. 1B, lane 1). As expected, translation initiation at the GUS start codon also produced a protein of 67 kDa (Fig. 1B, lane 7). This result showed that, in fact, the three hypothetical start codons were recognized, resulting in three different translation initiation products. Assuming that GUS translation initiation efficiency is similar in all situations, as can be envisioned by comparison of its intensity for all constructs, the intensity of the translational products from the first and second in-frame AUGs is expressed as regards the intensity of GUS initiation in each lane. Quantification analysis indicated that the 83 kDa protein is approximately three times more abundant than the 75 kDa protein (Fig. 1C, lane 1). When a similar experiment was performed using the THI64Δ1-GUS construct (Fig. 1A), in which the first AUG codon was converted to AUC (Ile), translation initiation at the first start codon was significantly reduced (Fig. 1B, lane 2). Furthermore, translation initiation efficiency at the second in-frame AUG codon (75 kDa) (Fig. 1B, lane 2) was increased significantly by nearly two-fold (Fig. 1C). On the other hand, the THI64Δ2-GUS plasmid, in which the second in-frame AUG codon was mutated to AUC (Fig. 1A), was able to efficiently synthesize only the 83 kDa protein (Fig. 1B, lane 3), in spite of an internal initiation being observed bearing the mutated site. This might be due to the presence of a close in-frame AUG codon (five codons downstream from the second putative in-frame initiation codon) or an artifact of the in vitro system.

The context surrounding the first AUG plays a major role in the regulation of translation initiation

An analysis of the translation efficiency of *thi1* mRNA revealed that translation occurs preferentially in the first AUG codon (Fig. 1B, lane 1; Fig. 1C, lane 1). On the other hand, the second in-frame AUG codon is inserted into a suboptimal context (**G**AGAUG**A**C). In order to figure out the context in the vicinity of both AUG start codons in *thi1* translation initiation, we performed site-directed mutagenesis creating more or less favorable contexts (Fig. 1A). A poor context around the first AUG start codon (**C**AAAUG**C**CU) (Fig. 1A) drastically reduced its translation efficiency (Fig. 1B, lane 4 and Fig. 1C, lane 4). In addition, translation initiation at the second in-frame AUG was increased. Interestingly, when an optimum context (**ACC**AUGGCG) was introduced into the vicinity of the second in-frame AUG (Fig. 1A), the efficiency of translation was not significantly affected (Fig. 1B, lane 5 and Fig. 1C, lane 5). When the translation initiation context was altered simultaneously (the suboptimal context **C**AAAUG**C**CU was introduced around the first start codon, and the more favorable context **ACC**AUG**G**CG was placed around the second in-frame start codon) (Fig. 1A), initiation at the upstream AUG was significantly reduced, whereas translation at the second site was increased by 50% (Fig. 1B, lane 6; Fig. 1C, lane 6), ensuring that in the absence of first AUG furtherance, second AUG improvement can be accomplished.

Dual targeting of thi1 to chloroplasts and mitochondria is determined by an alternative translation initiation mechanism

The occurrence of two translational products raises the possibility that the two isoforms are translated from *thi1* mRNA, using two different in-frame start codons. To explore this hypothesis, we analyzed the in vivo transport of both translational products using green fluorescence protein (GFP) gene constructs (Fig. 2). cDNA encoding the *thi1* N-terminal region (Met-1 to Pro-119) was fused to the 5′ end of the *gfp* gene (named pCambia-THI50-GFP) (Fig. 2a). To confirm that these two proteins were translated from two distinct in-frame AUG start codons, two other constructions were prepared, pCambia-THI50Δ1-GFP (Fig. 2b), in which Met-1 was converted to Ile-1, and pCambia-THI50Δ2-GFP (Fig. 2c), in which Met-70 was mutated to Ile-70. In addition, a construct starting from the second in-frame AUG (Met-70) was also obtained (pCambia-THImet2-GFP) (Fig. 2d). All the constructs were fused in-frame to the 5′ end of the *gfp* gene and placed under the control of the transcriptional 35S promoter from the cauliflower mosaic virus. These plasmids were introduced into tobacco leaf protoplasts by electroporation, and transient expression was analyzed by fluorescence microscopy. After 24 hours of culture, the THI50-GFP chimeric protein could be observed in both mitochondria and chloroplasts (Fig. 3A). Chloroplasts could be easily distinguished by their red autofluorescence, owing to their chlorophyll. Similar results were obtained with the SYCO-GFP construct (Fig. 3B), which has been shown to be dual targeted to both organelles (Peeters et al., 2000). With the THI50Δ1-GFP construct (first AUG converted to AUC) the fluorescence was restricted to mitochondria (Fig. 3C). The fluorescence pattern of β-GFP, which is known to be selectively delivered into mitochondria (Duby et al., 2001), was very similar to THI50Δ1-GFP (Fig. 3D). Additionally, in the tobacco protoplasts transfected with THImet2-GFP (second in-frame AUG was fused to GFP), green fluorescence was only seen in mitochondria (Fig. 3E). By contrast, with the THI50Δ2-GFP construct (in which the second in-frame AUG was converted to AUC), fluorescence was restricted to chloroplasts (Fig. 3F). These results are similar to the typical plastid targeting construct RecA-GFP (Fig. 3G) (Köhler et al., 1997). With the 35S-GFP construct (not fused to any of the *thi1* targeting sequences), GFP fluorescence was found in the cytosol (data not shown).

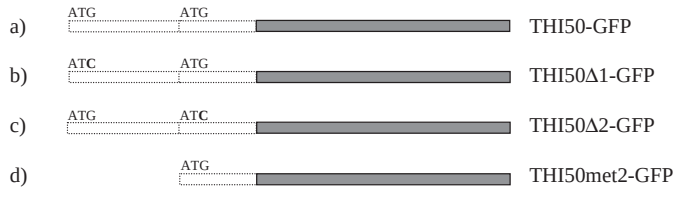


Fig. 2. Schematic representation of the GFP fusions used in transient expression experiments. (a) The THI50-GFP construct corresponds to the N-terminus of THI1 (first 105 amino acids) containing the entire chloroplast transit peptide followed by 50 amino acids of the mature THI1 (which includes the putative mitochondrial presequence), with both 5′ in-frame ATG codons. (b) THI50Δ1-GFP, the first ATG was mutated to ATC whereas the second in-frame start codon remained intact. (c) THI50Δ2-GFP, in which the second ATG was converted to ATC. (d) THImet2-GFP starts at the second ATG codon.

Structural analysis of the 5′ portion of the thi1 mRNA

To search for mechanisms that could explain the possible translation initiation at the second AUG, *thi1* mRNA was dissected. Two interesting features were observed that might be linked to translation initiation. First, *A. thaliana* wild-type *thi1* mRNA is predicted to form a relatively stable stem-loop structure between the two in-frame AUG codons, particularly in the region extending from the last base of codon 41 through the first base of the second in-frame AUG start codon (Fig. 4A). The stability of such a stable stem-loop structure at this position is interesting and may interfere in the translation efficiency. Second, we examined the sequences within the region flanked by the two in-frame AUGs start codons of *thi1*, which could function as a putative internal ribosome entry site (IRES). In previous studies, it has been shown that eukaryotic mRNAs contain short complementary matches to 18S rRNA, suggesting that ribosome recruitment at some cellular IRES might occur by base pairing between mRNA and 18S rRNA (Zhou et al., 2001; Fernandez et al., 2002). In fact, sequence comparison between *thi1* mRNA and 18S rRNA identified two complementary sequence matches (Fig. 4B). These matches contain stretches of 13 (nucleotides 62-74) and 12 (91-102 nt) nucleotides with 85% and 92% complementarity, respectively.

Discussion

In the present study, we have investigated the mechanism responsible for the dual targeting of *Arabidopsis thaliana* THI1

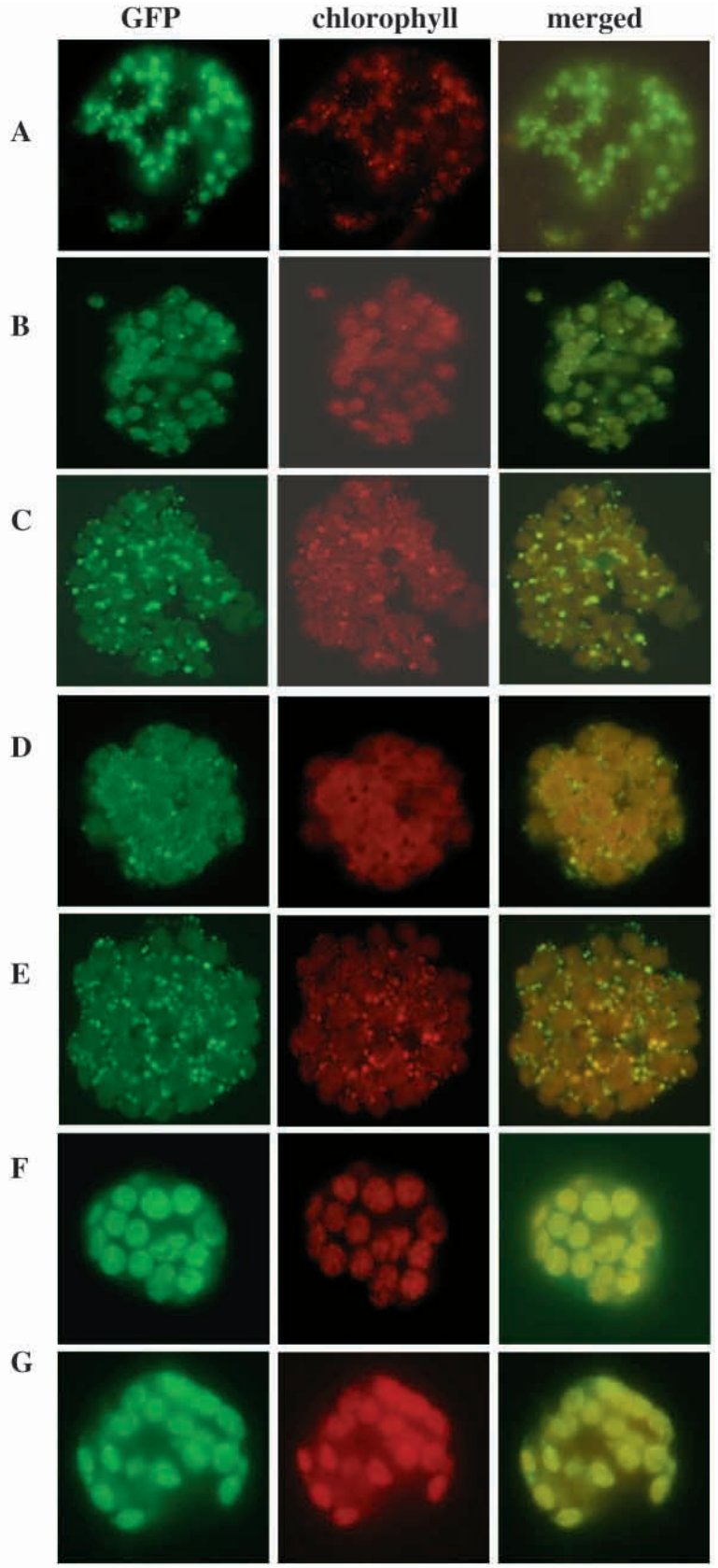


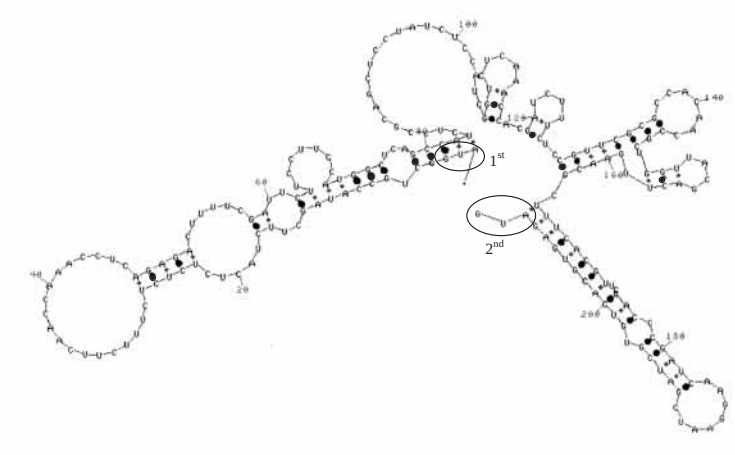
Fig. 3. Transient expression of GFP fusions in wild-type tobacco leaf protoplasts. (A) Expression of the THI50-GFP construct. (B) Dual targeting to plastid and mitochondria by SYCO-GFP (Peeters et al., 2000). (C) Expression of the THI50Δ1-GFP construct. (D) Mitochondrial targeting of GFP mediated by β-GFP construction (Duby et al., 2001). (E) Expression of the THImet2-GFP construct. (F) Expression of the THI50Δ2-GFP construct. (G) Plastid targeting of GFP by RecA-GFP (Köhler et al., 1997). The images were taken by epifluorescence microscope Axioplan 2. The fluorescence of GFP was observed at an excitation wavelength of 450-490 nm and emission wavelength of 520 nm (A-G, left panel). The autofluorescence of chloroplasts was observed at an excitation wavelength of 488 nm and emission wavelength of 664-696 nm (A-G, middle panel). The merged images of GFP and chlorophyll fluorescence are show in the right panel.

to mitochondria and chloroplasts, using a combination of in vitro transcription/translation experiments and in vivo GFP expression in tobacco leaf protoplasts.

The finding that *thi1* contains a single mRNA excludes the possibility of transcript heterogeneity, which could result in the production of proteins with variable N-termini. In vitro transcription/translation experiments showed that *thi1* produces two translational products with the expected size for translation initiation for the two in-frame AUG start codons. Interestingly, analysis of the vicinity of both start codons indicates that the sequence context around the first AUG (**A**AA<u>AUG</u>**G**C) is more favorable for translation initiation than the second (**G**AG<u>AUG</u>**A**C), as it corresponds to the reported dicot consensus (**A**/GAA<u>AUG</u>**G**C) (Joshi et al., 1997; Lukaszewicz et al., 2000). A purine, preferably A, at position –3 (three nucleotides before the AUG codon, which is numbered +1 to +3), and G at position +4, are the major determinants for translation initiation efficiency in higher plants (Kozak, 1991; Kozak, 1997). We reasoned that the observed low translation levels at the second AUG may, somehow, be related to a reduced demand for the shorter translational product inside the cell.

Since THI1 presents two targeting sequences in tandem, we hypothesized that translation initiation in the first AUG would deliver the protein to chloroplasts. In addition, the presence of a second in-frame AUG at the beginning of a sequence able to fold into an amphiphilic α-helix, and usually found in mitochondrially imported proteins (Neupert, 1997), suggested that this codon could be a site for translation initiation. The presence of both in-frame AUG codons fused to the 5′ end of GFP directed translocation of the reporter protein to mitochondria and chloroplasts of tobacco leaf protoplasts. Supporting evidence for this mechanism was obtained from the observation that mutation in the first AUG to AUC abolished importation of GFP to chloroplasts, and the fluorescent protein was found to be only associated with mitochondria. Thus, the second AUG can also be recognized in vivo by the plant translation machinery and the product indeed delivered to mitochondria. In accordance with this, the construct starting directly from the second in-frame AUG fused to GFP confirmed that this amphiphilic α-helix secondary structure functions as a mitochondrial presequence. Another

A)



B)

thi1 (62-74) CUUCCUUCCAUGG

18S rRNA (515-503) GACGGAAGGAACC

thi1 (91-102) GCUCCUAUCUCC

18S rRNA (482-472) CGAGGG-AGAGG

Fig. 4. Sequence analysis of the 5′-end of *thi1* mRNA. (A) Evidence for a stable stem-loop structure in a region between the two in-frame start codons (boxed). Predicted secondary structure of the 90 first codons of the wild-type *thi1* mRNA. Structure was generated for 28°C by using the mfold version 2.3 program (http://mfold2.wustl.edu/~mfold/rna/for1-2.3.cgi). Black dots indicate G:C base pairs and grey dots indicate weaker A:U or G:U base pairing. (B) Complementary sequence between *thi1* mRNA and 18S rRNA of *A. thaliana*. Vertical lines indicate base pairing and open circles represent G:U base pairing.

important result was obtained by the conversion of the second in-frame AUG to AUC: GFP was then found associated to the plastids and the targeting to mitochondria completely lost. The fact that GFP was targeted specifically to particular organelles is not direct proof of targeting, but does provide strong evidence that the different isoforms of THI1 are produced in vivo and that a translational initiation mechanism is involved in the dual targeting of this protein.

According to the scanning model, translation is initiated at the first AUG codon contained in a particular context (Kozak, 1991). Why would translation initiation occur at the second AUG of *thi1* mRNA if its context were not favored? It has been suggested that initiation sites in eukaryotic mRNAs are reached via a scanning mechanism that predicts that translation should start at the AUG codon nearest the 5′ end of the mRNA (Kozak, 1999). However, a recent survey on translation initiation in vertebrates indicates that translation initiation from downstream AUGs is more common than generally believed (Peri and Pandey, 2001). It is suggested that mechanisms such as leaky scanning, reinitiation or internal initiation of translation might have a greater role than previously reported (for a review, see Gray and Wickens, 1998). Translation at the second AUG, in which the surrounding sequence is not suitable for efficient initiation, might indicate a leaky scanning mechanism, where the ribosome does not always recognize the initial AUG. Alternatively, *thi1* mRNA contains short complementary matches to 18S rRNA, raising the possibility that ribosome recruitment at some cellular IRESes might occur by base pairing between *thi1* mRNA and 18S rRNA. Placing these observations together, translation of both in-frame start codons would involve a cap-dependent translation mechanism around the first AUG and a cap-independent translation process at the second start codon. Although speculative, these assumptions can be tested experimentally. Further studies are necessary to determine the mechanism for initiating translation of *Arabidopsis thaliana thi1* mRNA.

Although the function of THI1 in mitochondria and chloroplasts is not completely understood, it seems that THI1 requirement is increased in chloroplasts. In fact, thiamine biosynthesis has been associated to plastids, where it participates as a cofactor for the two enzyme complexes involved in the citric acid cycle, pyruvate dehydrogenase and α-ketoglutarate dehydrogenase (Belanger et al., 1995). The need for thiamine is thus essential for plant metabolism, which suggests that the enzymes involved in its biosynthetic pathway are required at high levels. On the other hand, it has been suggested that THI1 has a second role in protecting mitochondrial DNA from damage (Machado et al., 1996; Machado et al., 1997). One reasonable possibility is that THI1 participates in independent reactions in two organelles at different requirements. Evaluation of the precise role of THI1 in mitochondria and chloroplasts would certainly help in understanding the differential requirements of this protein and its effect on translation regulation.

The authors thank Marc Boutry for critical reading of the manuscript. We are very grateful to Ian Small and Marc Boutry for the generous gift of GFP-plasmids. We also thank Margarida L. R. Aguiar-Perecin for the working facilities with the epifluorescence microscope. This work was supported by Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP, São Paulo, Brazil) grant 98/11119-7. S.M.C. and D.D.L. were supported by graduate fellowships from FAPESP. M.C.S.F., C.F.M.M. and M.-A.V.S. are staff members of their respective departments and research fellows of CNPq (Brasília, Brazil).

References

Belanger, F. C., Leustek, T., Chu, B. and Kriz, A. L. (1995). Evidence for the thiamine biosynthetic pathway in higher-plant plastids and its developmental regulation. *Plant Mol. Biol.* **29**, 809-821.

Chabregas, S. M., Luche, D. D., Farias, L. P., Ribeiro, A. F., van Sluys, M.-A., Menck, C. F. M. and Silva-Filho, M. C. (2001). Dual targeting properties of the N-terminal signal sequence of *Arabidopsis thaliana* THI1 protein to mitochondria and chloroplasts. *Plant Mol. Biol.* **46**, 639-650.

Danpure, C. (1995). How can the products of a single gene be localized to more than one intracellular compartment? *Trends Cell Biol.* **5**, 230-238.

Duby, G., Oufattole, M. and Boutry, M. (2001). Hydrophobic residues within the predicted N-terminal amphiphilic α-helix of a plant mitochondrial targeting resequence play a major role in *in vivo* import. *Plant J.* **6**, 539-549.

Duchêne, A.-M., Peteres, N., Dietrich, A., Cosset, A., Small, I. D. and Wintz, H. (2001). Overlapping destinations for two dual targeted glycil-tRNA synthetases in *Arabidopsis thaliana* and *Phaseolus vulgaris*. *J. Biol. Chem.* **276**, 15275-15283.

Fernandez, J., Yaman, I., Merrick, W. C, Koromilas, A., Wek, R. C., Sood,

R., Hensold, J. and Hatzoglou, M. (2002). Regulation of internal ribosome entry site-mediated translation by eukaryotic initiation factor-2α phosphorylation and translation of a small upstream open reading frame. *J. Biol. Chem*, **277**, 2050-2058.

Franzén, L. G., Rochaix, J. D. and von Heijne, G. (1990). Chloroplast transit peptides from the green alga *Chlamydomonas reinhardtii* share features with both mitochondrial and higher plant chloroplast presequences. *FEBS Lett.* **260**, 165-168.

Frearson, E. M., Power, J. B. and Cocking, E. C. (1973). The isolation, culture and regeneration of Petunia leaf protoplasts. *Dev. Biol*. **33**, 130-137.

Gabriel, K., Buchanan, S. K. and Lithgow, T. (2001). The alpha and the beta: protein translocation across mitochondrial and plastid outer membranes. *Trends Biochem. Sci.* **26**, 36-40.

Gebhardt, J. S., Wadsworth, G. J. and Matthews, B. F. (1998). Characterization of a single soybean cDNA encoding cytosolic and glyoxysomal isozymes of aspartate aminotransferase. *Plant Mol. Biol.* **37**, 99-108.

Glaser, E., Sjöling, S., Tanudji, M. and Whelan, J. (1998). Mitochondrial protein import in plants. *Plant Mol. Biol.* **38**, 311-338.

Gray, N. K. and Wickens, M. (1998). Control of translation initiation in animals. *Annu. Rev. Cell Dev. Biol.* **14**, 399-458.

Hedtke, B., Börner, T. and Weihe, A. (2000). One polymerase serving two genomes. *EMBO Rep.* **1**, 435-440.

Hurt, E. C., Soltanifar, N., Goldschmidt-Clermont, M., Rochaix, J.-D. and Schatz, G. (1986). The clevable pre-sequence of an imported chloroplast protein directs attached polypeptides into yeast mitochondria. *EMBO J.* **5**, 1343-1350.

Joshi, C. P., Zhou, H., Huang, X. and Chiang, V. L. (1997). Context sequences of translation initiation codon in plants. *Plant Mol Biol*. **35**, 993-1001.

Kobayashi, Y., Dokiya, Y. and Sugita, M. (2001). Dual targeting of phage-type RNA polymerase to both mitochondria and plastids is due to alternative translation initiation in single transcripts. *Biochem. Biophys. Res. Commun.* **289**, 1106-1113.

Köhler, R. H., Zipfel, W. R., Webb, W. W. and Hanson, M. R. (1997). The green fluorescent protein as a marker to visualize plant mitochondria *in vivo*. *Plant J*. **11**, 613-621.

Kozak, M. (1991). Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem*. **266**, 19867-19870.

Kozak, M. (1997). Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6. *EMBO J.* **16**, 2482-2492.

Kozak, M. (1999). Initiation of translation in prokaryotes and eukaryotes. *Gene* **234**, 187-208.

Lukaszewicz, M., Feuermann, M., Jérouville, B., Stas, A. and Boutry, M. (2000). In vivo evaluation of the context sequence of the translation initiation codon in plants. *Plant Sci*. **154**, 89-98.

Macasev, D., Newbigin, E., Whelan, J. and Lithgow, T. (2000). How do plant mitochondria avoid importing chloroplast proteins? Components of the import apparatus Tom20 and Tom22 from *Arabidopsis* differ from their fungal counterparts. *Plant Physiol*. **123**, 811-816.

Machado, C. R., Costa de Oliveira, R. L., Boiteux, S., Praekelt, U. M., Meacock, P. A. and Menck, C. F. M. (1996). *Thi1*, a thiamine biosynthetic gene in *Arabidopsis thaliana*, complements defects in DNA repair. *Plant Mol. Biol*. **31**, 585-593.

Machado, C. R., Praekelt, U. M., Costa de Oliveira, R. L., Barbosa, A. C. C., Byrne, K. L., Meacock, P. A. and Menck, C. F. M. (1997). Dual role for the yeast THI4 gene in thiamine biosynthesis and DNA damage tolerance. *J. Mol. Biol*. **273**, 114-121.

Neupert, W. (1997). Protein import into mitochondria. *Annu. Rev. Biochem*. **66**, 863-917.

Peeters, N. M., Chapron, A., Giritch, A., Grandjean, O., Lancelin, D., Lhomme, T., Vivrel, A. and Small, I. (2000). Duplication and quadruplication of *Arabidopsis thaliana* cysteinyl- and asparaginyl-tRNA synthetase genes of organellar origin. *J. Mol. Evol.* **50**, 413-423.

Peeters, N. and Small, I. (2001). Dual targeting to mitochondria and chloroplasts. *Biochim. Biophys. Acta.* **1541**, 54-63.

Peri, S. and Pandey, A. (2001). A reassessment of the translation initiation codon in vertebrates. *Trends Genet*. **17**, 685-687.

Roberts, C. S., Rajagopal, S., Smith, L. A., Nguyen, T. A., Yang, W., Nugroho, S., Ravi, K. S., Cao, M.-L., Vijayachandra, K., Patell, V. et al. (1997). A comprehensive set of modular vectors for advanced manipulations and efficient transformation of plants by both *Agrobacterium* and direct DNA uptake methods. In *Rockefeller Foundation Meeting of the International Program on Rice Biotechnology*, 15-19. Mallaca, Malaysia.

Sambrook, J. and Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press.

Sass, E., Blachinsky, E., Karniely, S. and Pines, O. (2001). Mitochondrial and cytosolic isoforms of yeast fumarase are derivatives of a single translation product and have identical amino termini. *J. Biol. Chem*. **276**, 46111-46117.

Silva-Filho, M. C. (1999). Translocation of a reporter protein into mitochondria is mediated by a chloroplast transit peptide and follows a normal import route. *J. Plant Physiol.* **151**, 51-54.

Silva-Filho, M. C., Chaumont, F., Leterme, S. and Boutry, M. (1996). Mitochondrial and chloroplast targeting sequences in tandem modify protein import specificity in plant organelles. *Plant Mol. Biol*. **30**, 769-780.

Silva-Filho, M. C., Wieërs, M.-C., Flügge, U.-I., Chaumont, F. and Boutry, M. (1997). Different *in vitro* and *in vivo* targeting properties of the transit peptide of a chloroplast envelope inner membrane protein. *J. Biol. Chem*. **272**, 15264-15269.

Small, I., Wintz, H., Akashi, K. and Mireau, H. (1998). Two birds with one stone: genes that encode products targeted to two or more compartments. *Plant Mol. Biol*. **38**, 265-277.

Soll, J. and Tien, R. (1998). Protein translocation into and across the chloroplastic envelope membranes. *Plant Mol. Biol.* **38**, 191-207.

Watanabe, N., Che, F.-S., Iwano, M., Yakayama, S., Yoshida, S. and Isogai, A. (2001). Dual targeting of spinach protoporphyrinogen oxidase II to mitochondria and chloroplasts by alternative use of two in-frame initiation codons. *J. Biol. Chem*. **276**, 20474-20481.

Whelan, J., Knorpp, C. and Glaser, E. (1990). Sorting of precursor proteins between isolated spinach leaf mitochondria and chloroplasts. *Plant Mol. Biol*. **14**, 977-982.

Zhou, W., Edelman, G. E. and Mauro, V. P. (2001). Transcript leader regions of two *Saccharomyces cerevisiae* mRNAs contain internal ribosome entry sites that function in living cells. *Proc. Natl. Acad. Sci USA* **98**, 1531-1536.