

The Same Arabidopsis Gene Encodes Both Cytosolic and Mitochondrial Alanyl-tRNA Synthetases

Hakim Mireau, Dominique Lancelin, and Ian D. Small¹

Station de Génétique et d'Amélioration des Plantes, Institut National de la Recherche Agronomique, Route de St.-Cyr, Versailles F-78026 Cedex, France

In plants, all aminoacyl-tRNA synthetases are nuclearly encoded, despite the fact that their activities are required in the three protein-synthesizing cell compartments (cytosol, mitochondria, and chloroplasts). To investigate targeting of these enzymes, we cloned cDNAs encoding alanyl-tRNA synthetase (AlaRS) and the corresponding nuclear gene, *ALATS*, from Arabidopsis by using degenerate polymerase chain reaction primers based on highly conserved regions shared between known AlaRSs from other organisms. Analysis of the transcription of the gene showed the presence of two potential translation initiation codons in some *ALATS* mRNAs. Translation from the upstream AUG would generate an N-terminal extension with features characteristic of mitochondrial targeting peptides. A polyclonal antibody raised against part of the Arabidopsis AlaRS revealed that the Arabidopsis cytosolic and mitochondrial AlaRSs are immunologically similar, suggesting that both isoforms are encoded by the *ALATS* gene. In vitro experiments confirmed that two polypeptides can be translated from *ALATS* transcripts, with most ribosomes initiating on the downstream AUG to give the shorter polypeptide corresponding in size to the cytosolic enzyme. The ability of the presequence encoded between the two initiation codons to direct polypeptides to mitochondria was demonstrated by expression of fusion proteins in tobacco protoplasts and in yeast. We conclude that the *ALATS* gene encodes both the cytosolic and the mitochondrial forms of AlaRS, depending on which of the two AUG codons is used to initiate translation.

INTRODUCTION

The translation of a nucleotide sequence into an amino acid sequence is essentially assured by tRNAs. These molecular adapters must be doubly specific, both for their corresponding codon(s) and for the amino acid with which they are charged. Although codon recognition simply rests on base pairing with the anticodon, amino acid specificity is assured by the 20 aminoacyl-tRNA synthetases in a two-step reaction that involves the ATP-dependant activation of the amino acid followed by transfer of the activated amino acid to the 3' end of the tRNA (reviewed in Moras, 1992; Carter, 1993). The astonishing specificity of aminoacyl-tRNA synthetases for their cognate tRNAs and amino acids contributes to the high fidelity of the translation process. The manner by which aminoacyl-tRNA synthetases distinguish between different tRNAs despite their conserved secondary and tertiary structure is an area of considerable research (McClain, 1993; Saks et al., 1994). In most cases, tRNA recognition depends on only a small number of nucleotides (often in the anticodon or the acceptor stem). One of the most striking examples is the interaction between tRNA^{Ala} and alanyl-tRNA synthetase (AlaRS). AlaRS apparently primarily recognizes the conserved G₃:U₇₀ wobble base pair in the acceptor stem (Hou and Schimmel, 1988; Musier-

Forsyth et al., 1991; Gabriel et al., 1996). It has been shown that mutation of U₇₀ to C₇₀ abolishes aminoacylation of tRNA^{Ala} by AlaRS in a wide range of organisms (Hou and Schimmel, 1989), including plants (Carneiro et al., 1994). The simplicity and high conservation of this motif make it an ideal model for studying aminoacyl-tRNA synthetase-tRNA interactions.

In plant cells, aminoacyl-tRNA synthetases are required in the three protein-synthesizing compartments (cytosol, mitochondria, and chloroplasts). Because none of these enzymes has been shown to be encoded by the plant organellar genomes, it is likely that they are nuclearly encoded and imported from the cytosol. In general, the organellar enzymes have different substrate specificities than their cytosolic counterparts and resemble bacterial enzymes. For example, the plastid AlaRS can be clearly differentiated from the cytosolic enzyme by its ability to aminoacylate bacterial and plastid tRNA^{Ala}, which are poor substrates for the cytosolic enzyme (Steinmetz and Weil, 1986). The exceptions are the mitochondrial enzymes corresponding to imported tRNAs. In plant cells, at least one-third of the tRNAs involved in translation of mitochondrial mRNAs (including tRNA^{Ala}) are nuclearly encoded and imported from the cytosol (Dietrich et al., 1992). In the cases that have been investigated, the mitochondrial enzymes aminoacylating these tRNAs are indistinguishable from their cytosolic counterparts and are suspected of aiding in the import of tRNAs into mitochondria (Dietrich et al., 1992; Small et al., 1992).

¹ To whom correspondence should be addressed.

We have shown that mutation of U₇₀ to C₇₀ in Arabidopsis tRNA^{Ala} abolishes both its aminoacylation (Carneiro et al., 1994) and its import into plant mitochondria in vivo (A. Dietrich, L. Maréchal-Drouard, V. Carneiro, A. Cosset, I.D. Small, manuscript submitted). Similarly, the yeast tRNA^{Lys}(CUU), which is imported into mitochondria, passes through the protein import pathway, and the process requires both the cytosolic lysyl-tRNA synthetase and the precursor for the mitochondrial lysyl-tRNA synthetase (Tarassov and Entelis, 1992; Tarassov et al., 1995a, 1995b). To investigate some of these processes in more detail, we decided to clone the nuclear gene encoding the mitochondrial AlaRS from Arabidopsis.

RESULTS

The Cloning of Arabidopsis cDNAs Encoding AlaRS

The strategy used to clone Arabidopsis AlaRS cDNAs was based on successive polymerase chain reaction (PCR) am-

plifications that used degenerate primers deduced from regions highly conserved between the *Escherichia coli* (Putney et al., 1981) and silkworm (Chang and Dignam, 1990) AlaRSs. The substrate used was cDNA made from poly(A)⁺ RNA extracted from young Arabidopsis seedlings. The PCR procedure gave us a 196-bp probe that was used to screen an Arabidopsis cDNA library. Eventually, we obtained and sequenced 12 incomplete cDNA clones. Their sequences suggested that they were derived from the same Arabidopsis gene (all overlaps were 100% identical) and when combined gave us a 3147-bp cDNA. This cDNA sequence contained an open reading frame (ORF) that continued to the 5' terminus, suggesting that the reconstituted sequence was not full length. In fact, because the cDNA used to make the library was not tailed, it would have been virtually impossible to recover the 5' terminus from the library. There was an ATG codon close to the position expected for the initiation codon by comparison with AlaRS sequences from other organisms, which if used would give a polypeptide of 955 amino acids, as shown in Figure 1. The polypeptide potentially encoded by this ORF is slightly >50% identical to the cytosolic AlaRSs from silkworm (Chang and Dignam, 1990),

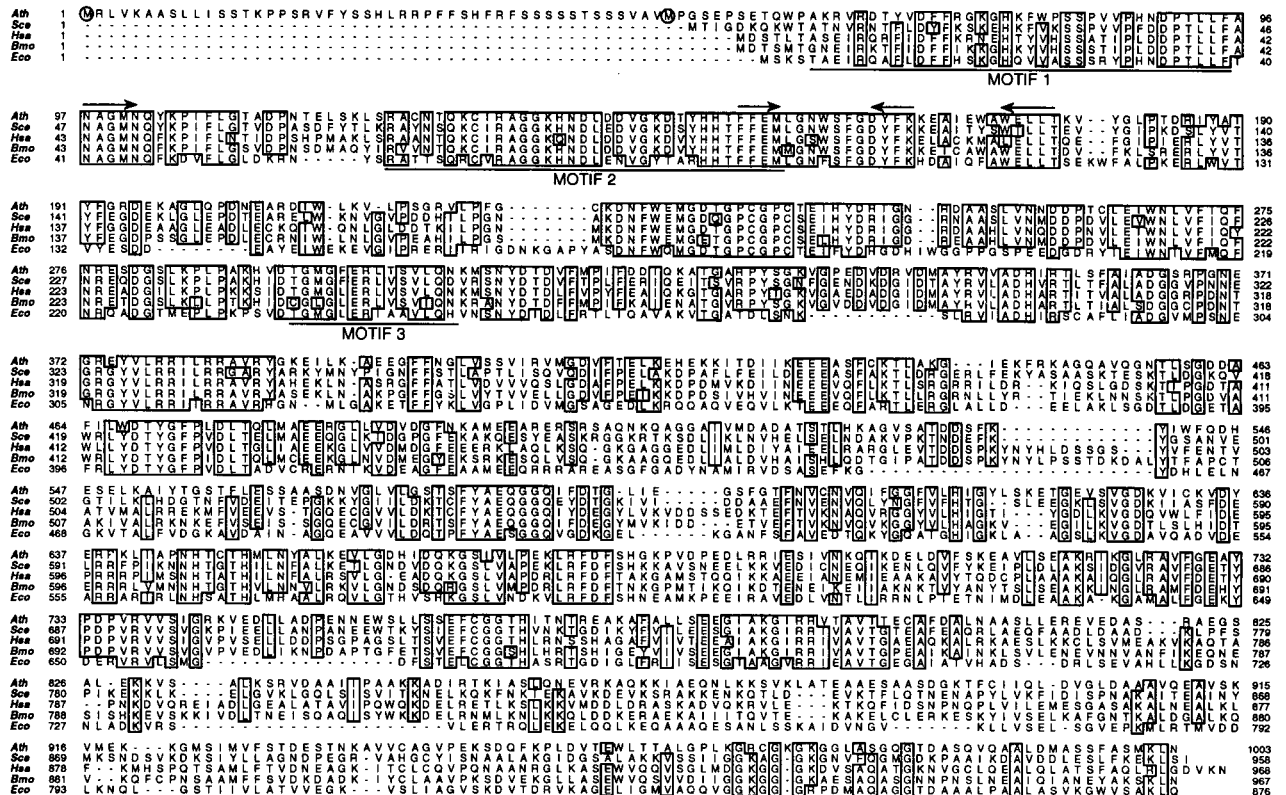


Figure 1. AlaRS Alignments.

A comparison of AlaRS amino acid sequences from Arabidopsis (*Ath*; EMBL accession number Z22673), baker's yeast (*Sc*; GenBank accession number U18672), humans (*Hsa*; GenBank accession number D32050), silkworm (*Bmo*; GenBank accession number M55993), and *E. coli* (*Eco*; GenBank accession number J01581, incorporating sequence corrections noted by Ribas de Pouplana et al. [1993]). The two possible initiator methionines (circled) correspond to polypeptides of 1003 and 955 amino acids, respectively. The regions used for designing the degenerate oligonucleotides used to amplify the Arabidopsis sequence are indicated by arrows. The three motifs that characterize AlaRS as a class II aminoacyl-tRNA synthetase are underlined with solid lines. Amino acids that are identical in four or more sequences are boxed.

yeast (Ripmaster et al., 1995), and humans (Shiba et al., 1995), but is only 41.7% identical to the *E. coli* AlaRS (Putney et al., 1981). In all cases, the similarity is higher in the N-terminal 400 to 500 amino acids that have been shown to contain the catalytic domain (Jasin et al., 1983; Buechter and Schimmel, 1993; Ribas de Pouplana et al., 1993). The high level of similarity assured us of the identity of the encoded enzyme as an AlaRS.

A Single Gene Potentially Could Encode the Cytosolic and the Mitochondrial Enzymes

We also attempted to clone the Arabidopsis genes encoding AlaRS. Because AlaRS is needed in the cytosol, plastids, and mitochondria, there are potentially three different nuclear genes encoding these three isozymes, and the biochemical data strongly suggest that at least the cytosolic and plastid enzymes are encoded by different genes (Steinmetz and Weil, 1986). However, tRNA^{Ala} is identical in the cytosolic and mitochondrial compartments of plant cells (Maréchal-Drouard et al., 1990), so AlaRS activity is likely to be similar or identical as well. Restriction map analysis of several of the genomic clones we obtained (Figure 2A and data not shown) showed that they came from the same region of the genome. This was confirmed by the DNA gel blot analysis in Figure 2B that resulted in, with the appropriate digestions, only one band of the expected size, using one of our cDNA clones as a probe. Under the hybridization and washing conditions used, we could not detect the gene encoding the chloroplast enzyme (probably ~40% identical to the amino acid sequence of the cytosolic enzyme), but we would have expected to detect two genes encoding the cytosolic and mitochondrial enzymes, respectively, if indeed the two isozymes were encoded by separate genes (assuming that their amino acid sequences were at least 90% identical). We concluded that the gene we had cloned had no other very close homolog in the Arabidopsis genome. We have named this gene *ALATS*, for alanyl-tRNA synthetase. Considering the similarity with the cytosolic silkworm, yeast, and human proteins, it probably encodes at least the cytosolic AlaRS.

If the *ALATS* gene encodes the mitochondrial enzyme, one would expect a mitochondrial targeting sequence to be encoded at the 5' end of the gene. Because the reading frame in our cDNA was open for >30 amino acids upstream of the first ATG, we sequenced the corresponding genomic region, which is shown in Figure 3A. We found another in-frame ATG (henceforth referred to as the upstream ATG) that could be the initiation codon for the translation of an ORF of 1003 amino acids containing an N-terminal extension (Figure 1). This potential initiation codon is 144 nucleotides upstream of the other ATG (henceforth referred to as the downstream ATG) and is preceded by potential TATA (Mukumoto et al., 1993) and CAAT motifs, suggesting that it could be transcribed. The amino acid composition and the predicted secondary structure of the peptide encoded between the two in-frame ATGs show several characteristics of mitochondrial presequences. There is a complete absence of acidic residues and an enrichment in Arg, Leu, and Ser residues. The first 15 amino acids are predicted

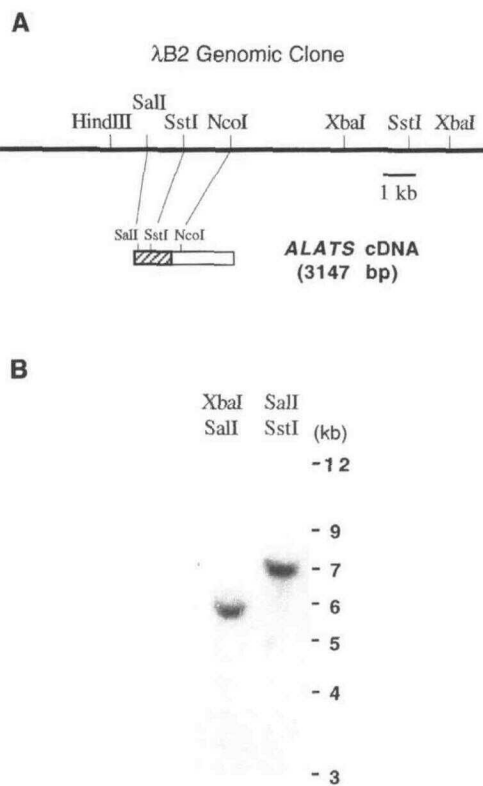


Figure 2. Analysis of the *ALATS* Gene.

(A) Restriction map of the 16-kb λ B2 genomic clone and the *ALATS* cDNA (showing selected sites). The NcoI-XbaI and XbaI-XbaI fragments were used as probes to complete the *ALATS* cDNA cloning. The cDNA region corresponding to the H17 cDNA clone is hatched.

(B) DNA gel blot of Arabidopsis DNA probed with *ALATS* cDNA. Three micrograms of total Arabidopsis DNA was doubly digested with the indicated enzymes. The filter was probed with an RNA probe synthesized from the H17 cDNA clone and washed under fairly stringent conditions (30 min at 65°C in 0.1 \times SSC [1 \times SSC is 0.15 M NaCl, 0.015 M sodium nitrate], 0.1% SDS). Molecular length markers are indicated at right.

to form an amphiphilic α -helix, and there is a short domain between amino acids 25 to 30, with a very high hydrophobic moment for angles between successive residues (θ) of ~75°, both features being frequently observed in a survey of 37 mitochondrial targeting peptides from different organisms (von Heijne et al., 1989).

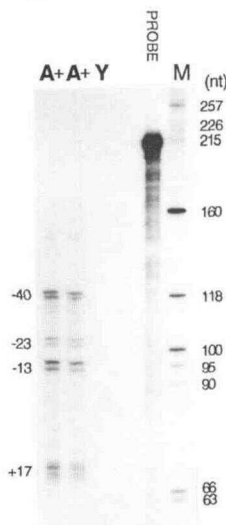
Transcription of the *ALATS* Gene

To verify that both ATG codons were transcribed, we determined the 5' termini of *ALATS* mRNAs (Figure 3). Preliminary indications were given by RNase protection experiments. We found complete protection of an antisense RNA probe extending from

A

...TTAATGTTATCCTTTGATGAACAAAATGTATTTTCCAAATTT
TACTTTTATTTTATTTTAAAGTATTTATCCCTTAAAATAATATATCT
Acil Apol
TGC CGAAAAGGAATCGGATTAATAATAATTTTAAACGGTAGAAAATA
GGATAGGAGGAGAGAGAGAAGGCTCTTTTATTGGTTTGAATGAGAT
slic2 oligo.
TAGTGAAGGCAGCTTCTCTCTGATTTCTTCAACCAAACCTCCGT
Alul Acil
CTAGGGTTTTTACTCTTCCCATCTCCGCCGCCTTTTTTCTCTC
slic1 oligo. Sall
ACTTCCGTTTTTCTTCTCTCTCGTCGACGCTCTCCCTCCGTCGCAG
TCATGCGGGTTCCGAACCATCGGAGACTCAGTGCCAGCAAAGA
GAGTCCGGGATACTTATGTCGATTTCTCAGAGGCAAAGGCCACAA
GTTCTGGCCACAAGCCCCGTTGTCTCCATAATGATCCTACTTTCT
TCTTCGCTAACGCTGGAATGAACCAAGTATAAGCCAATATTTTGG
GTACGCCGATCCAAATACTGAGCTTAGCAAGCTATCTCGGGCAT
GTAACACCCAAAAGTGTATTTCGTGCTGGTGGAAAGATAATGATTT
GGATGATGTCGGCAAAGATACGTACCATCACACCTTCTTTGAAAT
5580
GCTT...

B



C



Figure 3. Analysis of the 5' Extremities of *ALATS* mRNAs. **(A)** Partial sequence of an *ALATS* genomic clone showing the two in-frame ATG codons (indicated in boldface) in the 5' region of the gene. The CAAT- and TATA-like motifs of the probable promoter region are underlined. The mRNA extremities found by RNase protection **(B)** and

the Sall site to the Alul site just downstream of the upstream ATG, indicating that a major transcription initiation site is unlikely to be downstream of the Alul site (data not shown). A second RNase protection experiment done with an antisense probe extending between the two Acil sites bracketing the first ATG (Figure 3B) gave multiple protected fragments, suggesting that the 5' ends of *ALATS* mRNAs are heterogeneous. The estimated sizes of these protected fragments indicate three major populations of mRNAs beginning before the upstream ATG and another population where the 5' ends lie between the upstream ATG and the Alul site.

An independent and more precise determination of the 5' termini of *ALATS* mRNAs was made by using the single-strand ligation to single-strand cDNA (SLIC) technique (Dumas et al., 1991). As shown in Figure 3C, one major and several minor amplification products were obtained. The sequence of the major amplification product indicated a 5' terminus at position +17 with respect to the A residue of the upstream ATG (Figure 3A), but several other 5' termini were found that are distributed

PCR amplification of cDNA ends **(C)** are indicated by solid and open arrows, respectively. The size of the arrowheads represents approximately the frequency with which each 5' end was found with the two techniques (based on relative signal strengths for the RNase protection results and on 17 sequences for the PCR experiments). The oligonucleotides 5580, slic1, and slic2 used to identify the 5' mRNA extremities are indicated by arrows beneath the sequence. Homopurine and homopyrimidine sequences flanking the upstream ATG are boxed. The Acil sites delimiting the region used as an RNase protection probe in **(B)** are marked, as are the Apol and Sall sites used for many of the cloning experiments.

(B) Autoradiography of labeled antisense RNA probe fragments protected from RNase A digestion after hybridization to two independent preparations of Arabidopsis poly(A)⁺ RNA (A+) or yeast tRNA (Y). The products were run on an 8% polyacrylamide gel. The undigested probe is also shown (PROBE). Five micrograms of RNA was present in each sample, and RNase digestion was performed at 30°C for 30 min. Varying the quantity of RNA or the digestion conditions (from 15 min at 25°C to 30 min at 37°C) did not alter the banding pattern. At left are the estimated starting points of the different transcripts relative to the A residue of the first ATG codon, assuming that the RNA digestion products migrated 5 to 10% more slowly than did the labeled DNA fragments used as markers. Alul digest of pBluescript SK+ (M), with sizes in nucleotides (nt) indicated at right.

(C) PCR amplification of 5' ends of *ALATS* cDNAs. The first round of PCR used the upstream anchor primer and the 5580 oligonucleotide with Arabidopsis cDNA as template. One-fifth of the second-round PCR products were run on a 3% agarose gel and stained with ethidium bromide. Lane 1 contains no added template; lane 2, no reverse transcriptase added when the cDNA synthesis was done; lane 3, second-round PCR products obtained with the downstream anchor primer only; lane 4, second-round PCR products obtained with the downstream anchor primer and the *ALATS*-specific slic1 primer; lanes M, 1-kb DNA ladder (Life Technologies). The lengths of the marker fragments are indicated at right.

on both sides of the upstream ATG. The heterogeneity in apparent 5' termini is probably not due to the premature arrest of the reverse transcriptase, because reconstitution experiments under the same conditions using SLIC on *in vitro* transcripts of the cloned gene gave a single amplification product whose sequence showed the expected 5' terminus (data not shown). A second SLIC experiment using a primer overlapping the upstream ATG gave two major amplification products whose sequences indicated 5' termini at positions -23 and -40 with respect to the A residue of the upstream ATG (Figure 3A).

Although the RNase protection and SLIC experiments agreed very well as to the positions of the various 5' termini of the *ALATS* mRNAs, they disagreed as to their relative abundance. PCR-based methods favor the recovery of shorter fragments because they denature more easily and replicate faster. RNase protection tends to favor the detection of longer fragments because they hybridize better and contain more labeled UTP. Nevertheless, considering that the RNase protection pattern was unchanged by varying the quantity of RNA (data not shown), it seems likely that the probe was in sufficient excess to detect all *ALATS* mRNAs in the sample. In addition, the RNase protection pattern was unaltered by changing the digestion conditions (data not shown), suggesting that all the probe/mRNA hybrids were equally stable under the conditions employed in Figure 3B. For these reasons, the RNase protection results probably reflect the proportions of the different transcripts better than the nested SLIC procedure, which involved a total of 60 PCR cycles.

AlaRS Enzymes in the Cytosol and in Mitochondria Are Immunologically Similar

Rabbit polyclonal antibodies were raised against the C-terminal half of Arabidopsis AlaRS overexpressed in *E. coli*. This part of the polypeptide was chosen because it is less conserved than the N-terminal half and thus was more likely to generate specific antibodies. Figure 4A illustrates a protein gel blot analysis of Arabidopsis total and mitochondrial proteins by using this antibody preparation. There are similar immunoreactive polypeptides in both total and mitochondrial protein extracts that migrated with an apparent molecular mass of ~105 to 110 kD. The polypeptide in the mitochondrial preparation recognized by the AlaRS antibody is assumed to be inside the mitochondria because it was resistant to proteinase K treatment until the mitochondria were lysed with detergent. The identification and purity of the mitochondrial protein preparation were verified with appropriate controls (Figures 4B and 4C).

The molecular mass predicted from the amino acid sequence for the cytosolic enzyme is 105 kD, whereas that predicted for the presumed mitochondrial precursor is 110.5 kD; however, removal of the presequence by the mitochondrial processing peptidase could give a polypeptide similar in size to the cytosolic enzyme. The result strengthened our belief that the *ALATS* gene encodes both the cytosolic and mitochondrial enzymes.

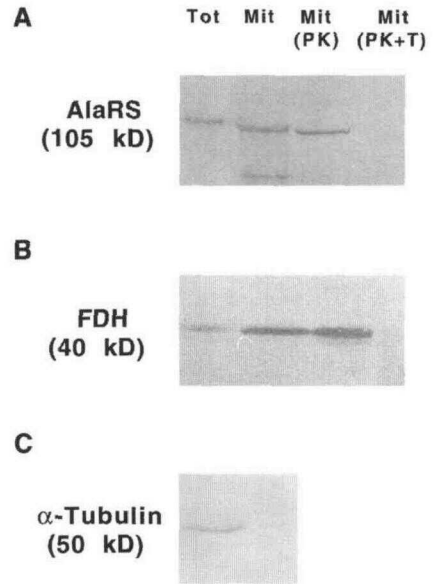


Figure 4. Protein Gel Blots Showing the Compartmentalization of AlaRS.

(A) Five micrograms of different Arabidopsis protein preparations was run on a 5% SDS-polyacrylamide gel, transferred to a nitrocellulose membrane, and probed with a polyclonal anti-Arabidopsis AlaRS antibody. Tot, total proteins; Mit, crude mitochondrial proteins; Mit(PK), proteins from mitochondria treated with proteinase K; Mit(PK+T), proteins from proteinase K-treated mitochondria that had been previously lysed with Triton X-100.

(B) Equivalently treated protein preparations were run on an 8% SDS-polyacrylamide gel and probed, after transfer to a nitrocellulose membrane, with polyclonal anti-formate dehydrogenase (FDH) antibody to reveal a typical soluble mitochondrial protein.

(C) A monoclonal anti- α -tubulin antibody was used as a control on total and crude mitochondrial protein preparations run on an 8% SDS-polyacrylamide gel. Only the total protein extract gave a signal, indicating little contamination of the mitochondrial preparation with cytosolic proteins. The sizes of the immunoreacting polypeptides were estimated from the migration of standard size markers on the same gels.

A Choice between Initiation Codons Occurs during Translation

Given the apparently high proportion of *ALATS* mRNAs that carry both potential initiation codons, we believed it necessary to verify which of these AUG codons was used preferentially in translation. The different AlaRS constructs diagrammed in Figure 5A were analyzed in an *in vitro* translation system. The 1WT2WT construct containing all of the elements present in the longest mRNAs, including all of the untranslated leader sequence and the two AUG codons, resulted in the synthesis of two polypeptides of >100 kD (Figures 5B and 5C). The smaller one, much more abundant than the other, comigrated with the polypeptide detected on protein gel blots of total Arabidopsis proteins (data not shown). Because this

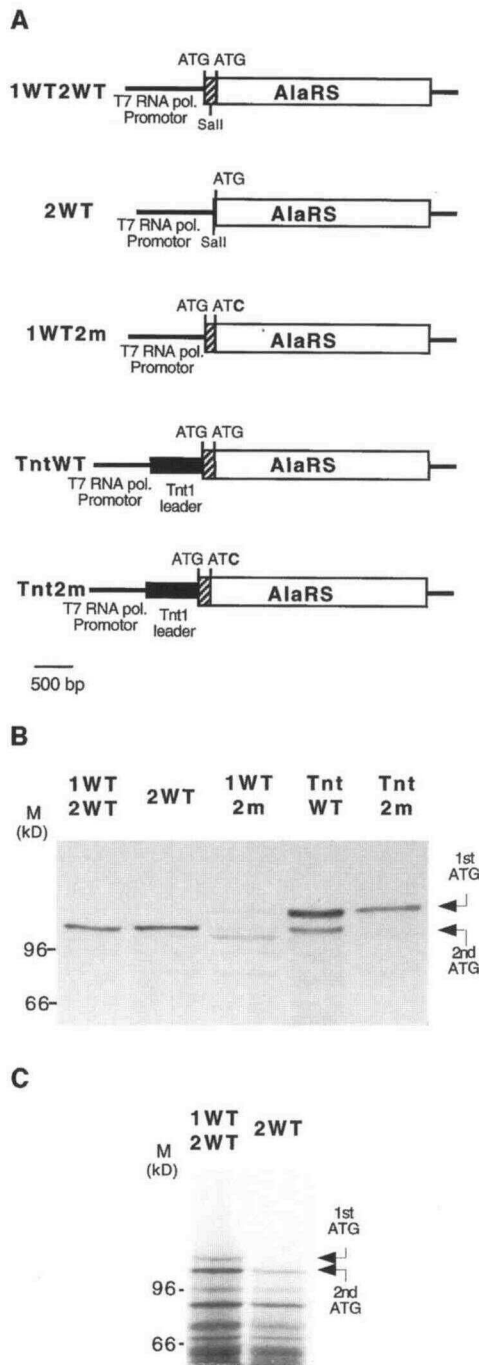


Figure 5. The First AUG Codon Is Inefficiently Used When Present on *ALATS* Transcripts.

(A) Constructs used for the *in vitro* translation experiments. The 1WT2WT construct contains both 5' ATG codons, whereas 2WT contains only the downstream ATG codon. In 1WT2m, the downstream ATG codon has been mutated to ATC. The TntWT construct has the Tnt1 transposon untranslated leader sequence in front of the upstream

major product was still produced after deletion of the upstream ATG (Figures 5B and 5C, construct 2WT) but was drastically reduced in abundance by mutation of the downstream ATG (Figure 5B, construct 1WT2m), it seems probable that it results in translation from the downstream AUG. The larger protein was only clearly visible when excess RNA was added to the translation reaction (Figure 5C, construct 1WT2WT) and is probably produced by translation from the upstream AUG, because it was not synthesized from the plasmid lacking the corresponding ATG (Figure 5C, construct 2WT). The apparent low efficiency of the first AUG codon to initiate translation led us to modify its environment by replacing the *ALATS* leader sequence (the untranslated 5' sequence) with the leader from the tobacco retrotransposon Tnt1 (Grandbastien et al., 1989). This leader has been shown to direct very efficient translation (N. Choisne, personal communication). The resulting Tnt1-AlaRS fusion showed greatly increased synthesis of the long form of AlaRS (Figure 5B, construct TntWT). Interestingly, there was little apparent reduction in the synthesis of the short form. Only the elimination of the downstream ATG by mutagenesis resulted in a construct (Figure 5B, construct Tnt2m) coding solely for the longest polypeptide.

The AlaRS Presequence Can Direct β -Glucuronidase and Yeast CoxIV into Mitochondria

To confirm the functional role of the putative AlaRS mitochondrial targeting sequence, we tested its ability to direct β -glucuronidase (GUS) into plant mitochondria. Figure 6A presents the different in-frame fusions between *ALATS* sequences and the *GUS* gene that were tested in transient expression experiments by electroporation of tobacco protoplasts. Total and mitochondrial GUS activities were measured in each case, and the relative specific activities showed whether a significant fraction of GUS enzymes had been directed to mitochondria, as shown in Figure 6B. The pre β -GUS plasmid, encoding an ATP synthase β subunit-GUS fusion (leading to

ATG codon. Tnt2m is identical to TntWT, except that the downstream ATG codon has been altered to ATC. pol., polymerase.

(B) Autoradiogram of ^{35}S -methionine-labeled *in vitro* translation products of the different constructs. One microgram of plasmid DNA was used for a coupled transcription/translation reaction. The molecular masses indicated at left show the migration of the corresponding protein size markers. The arrows at right indicate the products resulting from translation from the upstream ATG codon (1st ATG) or the downstream ATG codon (2nd ATG), respectively.

(C) Autoradiogram of ^{35}S -methionine-labeled *in vitro* translation products of 1WT2WT and 2WT. An excess of RNA (10 μg) was used for the translation reaction. The molecular masses shown at left indicate the migration of the corresponding protein size markers. The arrows on the right indicate the products resulting from translation from the upstream ATG codon (1st ATG) or the downstream ATG codon (2nd ATG), respectively.

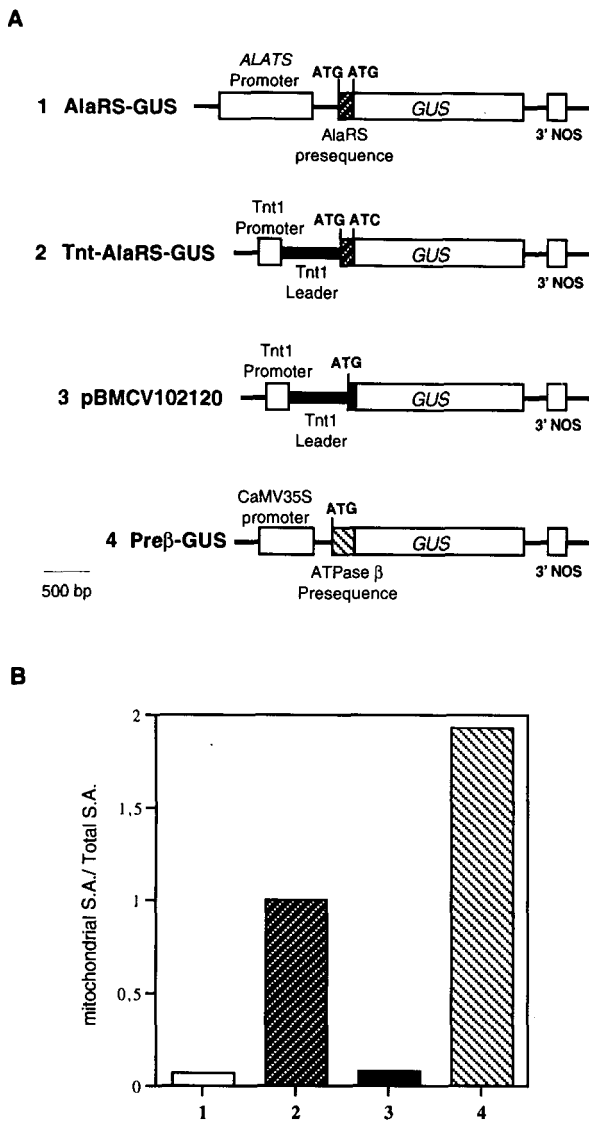


Figure 6. Import of GUS Fusion Proteins into Mitochondria.

(A) The two AlaRS–GUS fusions (AlaRS–GUS and Tnt–AlaRS–GUS), the Tnt1 polyprotein–GUS fusion (pBMCV102120), and the ATP synthase β subunit–GUS fusion (Preβ–GUS) are presented schematically. They were electroporated into tobacco protoplasts for transient expression experiments. NOS, nopaline synthase.

(B) Relative GUS activities measured 48 hr after electroporation. The ratio of the GUS specific activity found in mitochondrial proteins to the GUS specific activity found in total proteins is given for each plasmid. The experiment was repeated with similar results. The numbers below the figure refer to the different constructs presented in (A). S.A., specific activity.

mitochondrial GUS activity; Chaumont et al., 1994), and the plasmid pBMCV102120 (Pouteau et al., 1991), expressing a Tnt1 polyprotein–GUS fusion (leading to cytosolic GUS activity), were used as controls. The AlaRS–GUS construct, containing the ALATS promoter and both ATG codons, gave results similar to those when pBMCV102120 was used, indicating that virtually all of the enzyme produced is cytosolic. The Tnt–AlaRS–GUS construct was optimized for the synthesis of the putative mitochondrial precursor by replacing the ALATS promoter and 5' untranslated region with the Tnt1 promoter and leader and by altering the downstream ATG by mutagenesis. The results with this plasmid resembled those with preβ–GUS in that a significant level of activity was recovered in mitochondria. None of the plasmids gave rise to significant GUS activity in chloroplasts (data not shown).

A second functional test was done that relied on the previously demonstrated interchangeability of mitochondrial targeting sequences between plants and yeast (Schmitz and Lonsdale, 1989; Chaumont et al., 1990). We tested the ability of the AlaRS presequence to direct the CoxIV subunit of cytochrome c oxidase into yeast mitochondria in vivo. As shown in Figure 7A, different in-frame fusions were made between the 5' end of the ALATS gene and a COXIV gene lacking the region coding for its own targeting sequence; the resulting plasmids were used to complement the yeast strain WSR containing a disrupted chromosomal COXIV gene. Figure 7B shows the complementation observed by restoration of the respiratory activity on medium containing glycerol as energy source. The presence of the Tnt1 leader was necessary to observe growth on glycerol, implying that the upstream ATG was used too inefficiently in its usual context to provide sufficient mitochondrial protein to complement. Mutation of the upstream ATG abolished the ability of the AlaRS–CoxIV fusion to complement the mutant strain, confirming that this ATG is the initiation codon used to make the mitochondrial protein.

DISCUSSION

Sequence comparisons showed that AlaRS is a relatively well conserved enzyme, at least in the catalytic core region, and this facilitated the cloning of the Arabidopsis gene via a strategy based on sequence similarity. This high level of similarity leaves little doubt that the identified gene encodes an AlaRS, and virtually all of the residues and sequence motifs identified as important for activity of the *E. coli* enzyme are conserved (Buechter and Schimmel, 1993; Filley and Hill, 1993; Ribas de Pouplana et al., 1993; Davis et al., 1994; Lu and Hill, 1994; Shi et al., 1994).

The aminoacyl–tRNA synthetases have been divided into two distinct classes based on common architectural elements shared by the catalytic domains of the enzymes of each group (Eriani et al., 1990). The AlaRS enzyme belongs to the group

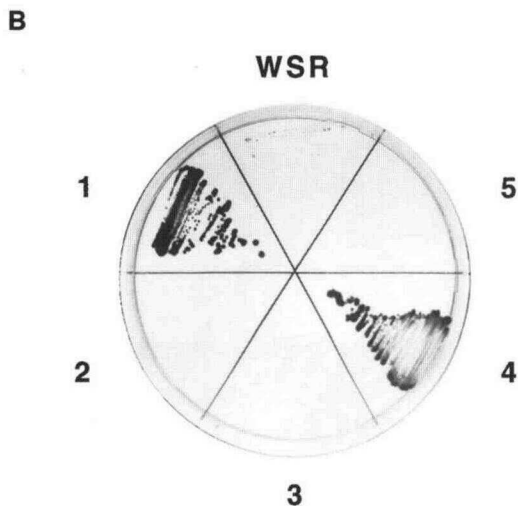
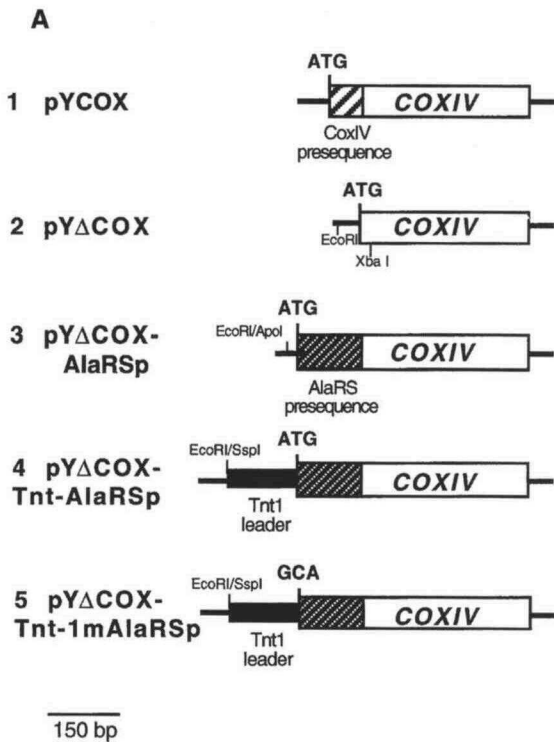


Figure 7. Complementation of a *CoxIV*-Deficient Yeast Strain by an AlaRS–*CoxIV* Fusion.

(A) The different constructs used in this study. The plasmid pYCOX carries the wild-type *COXIV* gene, including the region coding for the mitochondrial presequence, whereas in pY Δ COX this region has been deleted. In the pY Δ COX–AlaRSp construct, the region of the *ALATS* gene coding for the AlaRS presequence has been ligated in front of the partially deleted *COXIV* gene. The pY Δ COX–Tnt–AlaRSp construct contains the Tnt1 retrotransposon untranslated leader sequence in front

of class II aminoacyl–tRNA synthetases characterized by three highly degenerate sequence motifs—1, 2, and 3 (Figure 1). These motifs form part of the active site domain, which presents conserved secondary and tertiary structures consisting of seven or eight antiparallel β -strands and three α -helices. Study of the crystal structures of two other class II enzymes, yeast AspRS (Cavarelli et al., 1993, 1994) and *Thermus thermophilus* SerRS (Belrhali et al., 1994; Biou et al., 1994), has shown that motif 1 and the sequences immediately following it are involved in binding the tRNA acceptor stem, whereas motifs 2 and 3 are involved in ATP binding and catalysis. Originally, neither motif 1 nor 2 could be identified for *E. coli* AlaRS (Eriani et al., 1990), although a possible motif 2 was subsequently suggested (Cusack et al., 1991). Recent alignments of the Arabidopsis AlaRS sequence, along with other known AlaRSs, helped with identification of a potential motif 1 (Ribas de Pouplana et al., 1993) and realignment of motif 2 (Davis et al., 1994).

We have demonstrated that the Arabidopsis *ALATS* gene is a bifunctional gene coding for both cytosolic and mitochondrial AlaRS enzymes. In a similar fashion, the Arabidopsis glutathione reductase gene encodes both chloroplastic and mitochondrial isoforms as well as possibly a cytosolic isoform (Creissen et al., 1995). These genes belong to what is probably a fairly large class of genes encoding isozymes performing analogous functions in more than one cellular location (Danpure, 1995). Genes of this type that have been studied are usually characterized by the presence of more than one ATG codon in their 5' region, with the multiple enzyme destinations due to heterogeneous N-terminal ends, depending on which of the AUG codons is used to initiate translation. Often, such genes have alternative transcription start sites, leading to the synthesis of two mRNAs that differ by the presence or absence of the upstream region coding for a mitochondrial presequence. For example, the genes coding for the yeast valyl–tRNA (Chatton et al., 1988) and histidyl–tRNA (Natsoulis et al., 1986) synthetases and probably *Caenorhabditis elegans* histidyl–tRNA synthetase (Amaar and Baillie, 1993) are expressed in this fashion.

A different type of regulation has been described for the yeast genes encoding isopentenyl pyrophosphate:tRNA isopentenyltransferase (*MOD5*) and ATP(CTP):tRNA nucleotidyltransferase (*CCA1*) (Martin and Hopper, 1994) and the rat fumarase gene (Suzuki et al., 1992). In the case of *MOD5*, two transcripts are synthesized, and both contain two in-frame AUG codons in the 5' region. It has been demonstrated that alter-

of the *ALATS* segment. The pY Δ COX–Tnt–1mAlaRSp construct is identical to pY Δ COX–Tnt–AlaRSp, except that the AlaRS initiation codon has been altered from ATG to GCA.

(B) Restoration of respiration in the yeast strain WSR. Each of the constructs described in **(A)** was used to transform the strain WSR (containing a disrupted endogenous *COXIV*), and the transformed cells were streaked on glycerol medium. The numbers correspond to the constructs described in **(A)**.

ing the nucleotides surrounding the upstream AUG to match the yeast consensus sequence or creating a longer mRNA leader increases expression at the first AUG and decreases expression at the second AUG (Slusher et al., 1991). The *CCA1* gene and the rat fumarase genes also have heterogeneous transcription start sites leading to mRNAs carrying two or three in-frame 5' AUG codons (Suzuki et al., 1992; Wolfe et al., 1994). For both genes, the most active AUG in terms of translation initiation appears to be the downstream one, despite the presence of one or two upstream AUG codons. Thus, for *MOD5*, *CCA1*, and the rat fumarase gene, the choice of cytosolic or mitochondrial location is not made solely on the basis of the transcription start point but also presumably involves AUG selection by the ribosomes. For the products of *MOD5*, *CCA1* (Martin and Hopper, 1994), and the yeast fumarase gene (Stein et al., 1994), targeting by their presequences is inefficient or incomplete. Uniquely, all yeast fumarase precursor molecules are cleaved by the mitochondrial processing peptidase, but 90% of the cleaved polypeptides are released back into the cytosol, with only 10% imported into the mitochondrial matrix (Stein et al., 1994).

Our results do not conclusively show which of these potential mechanisms is responsible for the partitioning of Arabidopsis AlaRS between the cytosolic and mitochondrial compartments. It is clear that translation from the upstream AUG codon can give rise to mitochondrial protein, but we cannot be sure that all the polypeptides translated from this AUG end up in mitochondria. Both the RNase protection and the SLIC experiments indicate that several 5' termini exist for *ALATS* transcripts, and this heterogeneity can explain at least in part the generation of two polypeptides. However, even in the SLIC experiments, more transcripts contain the upstream AUG initiation codon than would be expected if the most 5' AUG codon was always used to initiate translation and if translation from the upstream AUG only gave the mitochondrial protein. The RNase protection results suggest that more than half of the *ALATS* mRNAs contain the upstream AUG codon.

Our in vitro translation results showed that a single population of *ALATS* transcripts containing both potential initiation codons can give rise to two proteins in the expected relative amounts (<5% mitochondrial precursor and the rest destined for the cytosol). It is thus possible that the dual targeting of AlaRS in plant cells is achieved primarily via AUG selection by ribosomes. If the standard scanning model applies to the *ALATS* mRNA, the upstream AUG must be extraordinarily inefficient. The preference for the downstream AUG was observed in three different systems (rabbit reticulocyte lysate, yeast cells, and tobacco protoplasts), which could be due to *cis* elements in the mRNA itself. A close examination of the sequence suggests two reasons for this. First, as described for the AUGs that are bypassed on the *MOD5* and *CCA1* mRNAs, the upstream AUG is placed in a poor context (UGAAUGAG) in comparison with the plant consensus sequence AC/A₃C/A₃AUGGC (Cavener and Stuart, 1991). The absence of the highly conserved purine at position -3 and the

guanosine at +4 predicts it to be a weak initiation codon in eukaryotes (Kozak, 1986). The downstream AUG is in a better context with a purine at position -3 and cytosine at +5 (GUCAUGCC). Second, there are partially complementary homopurine and homopyrimidine sequences on either side of the first AUG (Figure 3), which could form a relatively stable secondary structure blocking access to the AUG. The substantially increased expression from the upstream AUG in the presence of the Tnt1 leader (with adenine at -3) demonstrates clearly the inhibitory effect of the sequence that normally lies upstream of the first AUG, although it does not indicate which of the two suggested mechanisms is the most important.

METHODS

Except where stated otherwise, standard molecular biological techniques were performed using the protocols described by Ausubel et al. (1990). Full details of the procedures used and the construction of the various plasmids described are available on request. Oligonucleotides were synthesized with a 381A DNA synthesizer (Applied Biosystems, Foster City, CA) or bought from Eurogentec (Liège, Belgium), Oligo Express (Montreuil, France), or Genosys (Cambridge, UK). Sequencing was performed with fluorophore-labeled primers in a semi-automated sequencing system (model 370A; Applied Biosystems).

Poly(A)⁺ RNA Extraction and cDNA Preparation

Total RNA from *Arabidopsis thaliana* seedlings or inflorescences was isolated by the guanidine hydrochloride method (Logemann et al., 1987) and purified on a CsCl cushion. The poly(A)⁺ RNA fraction was recovered by using oligo(dT)-cellulose spun columns provided in an mRNA purification kit (Pharmacia Biotech, Uppsala, Sweden). For cDNA synthesis, 2 µg of poly(A)⁺ RNA was denatured at 65°C for 5 min and then placed on ice. The cDNA synthesis reaction was performed in 50 mM Tris-HCl, pH 8.3, 75 mM KCl, 3 mM MgCl₂, 10 mM DTT, 0.5 mM deoxynucleotides and in the presence of five units of RNase inhibitor (Life Technologies, Cergy Pontoise, France) and 200 units of Superscript murine leukemia virus reverse transcriptase (Life Technologies) for 2 hr at 42°C. Poly(A)⁺ RNA was then hydrolyzed with 0.3 N NaOH at 50°C for 15 min. The solution was neutralized with 0.3 N acetic acid, and the recovered cDNA was ethanol precipitated.

Cloning of AlaRS cDNAs and the AlaRS Gene

A 26-bp sequence specific to Arabidopsis cDNA was obtained by two successive polymerase chain reaction (PCR) amplifications using four nested degenerate oligonucleotides deduced from four regions conserved between alanyl-tRNA synthetases (AlaRS) from *Escherichia coli* (Putney et al., 1981) and silkworm (Chang and Dignam, 1990). The first PCR reaction was performed on cDNA from in vitro-grown Arabidopsis seedlings by using the two outer primers: 5'-AAGCNG-GNATGAA-3' (where N is A or C or G or T) and 5'-GTNARNARY-TCCCA-3' (where R is A or G; Y is C or T) deduced from the highly conserved NAGMN and WELLT regions of AlaRS (Figure 1). The PCR conditions were 30 cycles of 94°C for 1 min, 45°C for 1 min, and 72°C

for 30 sec plus 1 sec per cycle. The second PCR used 1 μ L of the first PCR products as substrate. This amplification was performed with the two inner primers, 5'-GATCAAGCTT^YTYGARATG-3' and 5'-GATCAGATCTT^YTRAARTATC-3', deduced from the two conserved regions FFEM and DYFK (Figure 1), respectively. The primers contained HindIII and BglII restriction sites (underlined) to facilitate the cloning of the PCR products. PCR conditions were five cycles of 94°C for 1 min, 37°C for 1 min, and 72°C for 20 sec followed by 30 cycles of 94°C for 1 min, 50°C for 1 min, and 72°C for 25 sec plus 1 sec per cycle. The 61-bp major amplification product was cloned, and its sequence was used to synthesize the oligonucleotide 5'-TCTCAAAGGACCAGTTACCAAGCAT-3', specific to the Arabidopsis cDNA, which was used with the first outer oligonucleotide (see above) to amplify a larger 198-bp fragment from the Arabidopsis AlaRS cDNA.

This fragment was cloned into pBluescript SK+ (Stratagene, La Jolla, CA), sequenced, and then transcribed in the presence of digoxigenin-UTP (Boehringer Mannheim France, Meylan, France). The probe obtained was used to screen 10^5 λ gt10 clones from an Arabidopsis ecotype Columbia cDNA library (Clontech, Palo Alto, CA) plated according to the manufacturer's instructions. Five independent, incomplete cDNA clones were obtained; their inserts were PCR amplified by using primers flanking the EcoRI cloning site and cloned into pBluescript SK+. Because none of them covered the entire cDNA sequence, the longest one, named H17 (Figure 2), was used as a probe to screen a genomic library (in EMBL-3 SP6/T7; Clontech). A dozen hybridizing genomic clones were purified and characterized by restriction digestions. Subsequently, two overlapping clones were partially sequenced. The NcoI-XbaI and the XbaI-XbaI fragments of the clone λ B2 (Figure 2A), containing the 3' end of the gene, were used to rescreen the cDNA library. Eventually, 12 recovered cDNA clones were sequenced, using single-stranded DNA prepared from subclones obtained from directed exonuclease III deletions.

Isolation of Total DNA and DNA Gel Blot Hybridization Analysis

Total DNA of greenhouse-grown plants of Arabidopsis was prepared by the method of Dellaporta et al. (1983) and further purified on CsCl gradients. For DNA gel blot analysis, 3 μ g of total DNA was digested with appropriate restriction enzymes, run on a 0.7% agarose gel, and transferred to Hybond-N nylon membranes (Amersham International, Little Chalfont, UK). RNA probes were labeled with α -³²P-UTP (Amersham) by T7 RNA polymerase transcription of the H17 cDNA clone.

Transcript Mapping

RNase protection experiments were done according to Bordonaro et al. (1994), except that the RNA probes were purified on a 6% polyacrylamide gel before use. The probes were recovered by overnight diffusion at room temperature in 200 μ L of 1.875 M ammonium acetate and 1% SDS. After precipitation, one-twentieth of the probe was used for overnight hybridization at 45°C with 5 μ g of Arabidopsis poly(A)⁺ RNA or yeast tRNA. The 5' ends of the Arabidopsis AlaRS transcripts were more precisely determined with the 5'-AmpliFINDER RACE kit (Clontech). cDNA enriched in AlaRS sequences was synthesized by using the 5580 AlaRS oligonucleotide 5'-GTACGTATCTTTGCC-3' (Figure 3A) to prime reverse transcription of poly(A)⁺ Arabidopsis RNA extracted from inflorescences. Two consecutive PCR amplifications were performed to obtain specific amplification products, the first with the anchor primer provided by the kit and the 5580

primer, the second with another anchor primer overlapping the first one (5'-TCCAGAATCGATAGTGAATTCGTG-3') and either of two AlaRS-specific primers close to the presumed 5' end of the mRNAs: the slic1 primer (5'-ACGTGGATCCGAAGAGTAAAAACCCTAGACGGAGGTTT-G-3') or the slic2 primer (5'-TCTCATCAAACCAATAAAAG-3') (Figure 3A). PCR reactions were performed according to the supplier's recommendations, and the PCR products obtained were cloned in pBluescript SK+ and sequenced.

Mitochondrial Purification

Approximately 20 g of Arabidopsis flowers, buds, and siliques were ground in a mortar with acid-washed sand and 5 mL of prechilled mitochondrial extraction buffer (0.4 M sucrose, 50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 5 mM β -mercaptoethanol, 0.5% BSA). The volume was then raised to 500 mL with extraction buffer, and the homogenate was filtered through four layers of Miracloth (Calbiochem, La Jolla, CA). All of the subsequent steps were performed at 4°C. The filtrate was centrifuged at 1500g for 10 min, and the pellet was discarded. The supernatant was recentrifuged at 3000g and then at 3500g (each spin was 10 min), and the pellets were discarded. The mitochondria in the supernatant were pelleted at 23,000g for 20 min and resuspended with a brush in 1 mL of resuspension buffer (0.4 M mannitol, 50 mM Tris-HCl, pH 7.5, 0.5% BSA). The organelles were finally layered on three-step Percoll (Pharmacia Biotech) gradients consisting of 4 mL of 50%, 5 mL of 25%, and 3.5 mL of 14% Percoll in 0.25 M sucrose, 50 mM Tris-HCl, pH 7.5, and spun at 17,000g for 20 min. The mitochondria were collected from the 50%:25% interface, diluted approximately six times with resuspension buffer, and pelleted at 17,000g for 20 min. For proteinase K treatment, the mitochondrial pellet was resuspended in 200 μ L of resuspension buffer devoid of BSA. The protein concentration was determined according to Bradford (1976). Four hundred micrograms of proteins was subjected to proteinase K digestion for 15 min on ice in the presence or absence of 0.5% Triton X-100. Either in the presence or absence of detergent, the formate dehydrogenase used as a mitochondrial marker protein proved to be much more resistant than AlaRS to the protease, presumably because of its much smaller size (40 kD) and greater abundance (Colas des Francs-Small et al., 1993). In the experiments shown in Figure 4, the minimal amounts of proteinase K needed to give complete digestion of the test polypeptide in the presence of Triton X-100 were added to each sample (7.5 μ g for AlaRS, 75 μ g for formate dehydrogenase). The proteinase K was inhibited by the addition of 2 mM phenylmethylsulfonyl fluoride (PMSF), and for the samples treated without Triton X-100, mitochondria were reisolated by pelleting through a 2-mL sucrose cushion (0.25 M sucrose, 50 mM Tris, pH 7.5, 1 mM PMSF) at 17,000g for 10 min.

Antibodies and Protein Gel Blot Analysis

A 1259-bp SspI-RsaI fragment from the 3' end of the Arabidopsis AlaRS cDNA (positions 1858 to 3116 of the cDNA sequence) was cloned in-frame at the filled-in XhoI site of the vector pET-14b (Novagen, Madison, WI). The in-frame fusion was checked by sequencing the junction. The peptide was overexpressed in the strain BL21(DE3)pLysS, according to Novagen's instructions, and purified on Ni-NTA resin (Qiagen Inc., Chatsworth, CA). Approximately 1 mg of purified protein in complete Freund's adjuvant was injected subcutaneously into a New Zealand male rabbit, and three more similar injections were done in incomplete Freund's adjuvant at 2-week intervals. Blood was collected 7 days after the last injection.

Total Arabidopsis proteins were prepared according to Damerval et al. (1986). For protein gel blots, 5 µg of each protein sample was run on 5% SDS-PAGE gels and transferred to Hybond-C membranes (Amersham). A 1:4000 anti-AlaRS antiserum dilution gave a strong, specific signal at the expected size, whereas preimmune serum gave no significant signal under the same conditions. The monoclonal anti- α -tubulin antibody (clone MAS 077B; Sera-Lab, Crawley Down, UK; used at 1:300 dilution) was kindly given by Jan Traas (Institut National de la Recherche Agronomique, Versailles, France), to check for contamination of mitochondrial protein preparations with cytosolic polypeptides. The polyclonal anti-formate dehydrogenase antibody (used at 1:4000 dilution) was kindly given by Catherine Colas des Francs-Small (Institut de Biotechnologie des Plantes, Orsay, France).

In Vitro Transcription/Translation of AlaRS Constructs

The whole of the cytosolic AlaRS coding sequence was ligated into pBluescript SK+ behind 1 kb of *ALATS* promoter region comprising the HindIII-Sall fragment isolated from the λ B2 genomic clone (Figure 2A) containing the upstream ATG. This construct was named P-tAlaRS. To make a plasmid suitable for in vitro transcription, P-tAlaRS was cut with Sall and religated to remove the *ALATS* promoter and the upstream ATG. This brought the downstream ATG close to the T7 RNA polymerase promoter, forming the 2WT construct. To bring back the upstream ATG, an Apol-SstI fragment of P-tAlaRS containing both ATGs was ligated to the SstI-BamHI fragment from the 2WT plasmid (containing the rest of the cDNA sequence) into pBluescript SK+ (cut with EcoRI and BamHI), resulting in the construct named 1WT2WT. To create the plasmid 1WT2m, the downstream ATG was mutated to ATC with the oligonucleotide 5'-TCGCAGTCATCCCGGTTCCG-3' (mutation underlined). The Tnt1 leader-AlaRS fusion (TntWT) was performed by PCR, using the overlapping primer 5'-TGGTGAAAATAAAAATGAGATTAGTGAAGGCA-3' (ATG initiation codon underlined). The fusion product was digested by HindIII and Sall, and ligated with the Sall-BamHI fragment containing the cytosolic AlaRS coding sequence into pBluescript SK+ cut with HindIII and BamHI. The introduction of the mutation at the downstream ATG was done by substituting the Sall-BamHI *ALATS* cDNA fragment from the previous fusion by the same insert isolated from the 1WT2m construct. All of the constructs were checked by sequencing the relevant portions.

In vitro-coupled transcription and translation were performed using the TNT Coupled Rabbit Reticulocyte Lysate System (Promega, Madison, WI), according to manufacturer's recommendations. In vitro transcriptions were performed using 3 to 5 µg of BamHI-linearized plasmids in RiboMAX buffer (Promega) with 50 units of T7 RNA polymerase (Life Technologies) for 2 hr at 37°C (another 50 units of T7 RNA polymerase were added after the first hour). The DNA was then digested with 27 units of RNase-free DNaseI (Boehringer Mannheim). After phenol-chloroform extraction, the RNAs were ethanol precipitated. Ten micrograms of these RNAs was translated in vitro with 17.5 µL of rabbit reticulocyte lysate (Promega), according to the manufacturer's instructions. One-tenth of the different translation products were run on 5% SDS-PAGE gels stained with Coomassie Brilliant Blue R 250 and dried before exposure.

AlaRS and β -Glucuronidase Fusions and β -Glucuronidase Assays

The AlaRS- β -glucuronidase (AlaRS-GUS) construct contains the *ALATS* promoter and both potential initiation codons fused to the *GUS*

coding sequence. The downstream *ALATS* ATG is placed exactly at the usual position of the *GUS* initiation codon. The fusion was performed by PCR (details available on request). The Tnt1-AlaRS presequence-GUS fusion was made by ligating a HindIII-Sall fragment from the Tnt2m construct (see above) containing the Tnt1 leader and the AlaRS presequence into the AlaRS-GUS plasmid at the same sites. To bring this fusion under the control of the protoplast-inducible Tnt1 promoter, a HindIII-BclI fragment containing the Tnt1 leader, the AlaRS presequence, and the start of the *GUS* coding sequence was cloned into the pBMCV102120 vector (Pouteau et al., 1991) at the same sites. All the constructs were finally checked by sequencing.

A plasmid expressing a fusion between the presequence of the β subunit of the mitochondrial ATP synthase and *GUS* and shown to direct *GUS* protein into plant mitochondria (Chaumont et al., 1994) was kindly provided by Marc Boutry (Université Catholique de Louvain, Belgium) and used as a control.

Tobacco leaf protoplast preparation, protoplast electroporation, total protein extracts, and *GUS* activity measurements were done as previously described (Carneiro et al., 1993). Purification of tobacco protoplast mitochondria was principally done as described above for Arabidopsis mitochondria, except that extraction was done in 60 mL of extraction buffer and the Percoll gradients were spun in 2-mL tubes containing 450 µL of 50%, 700 µL of 25%, and 500 µL of 14% Percoll. Recovered mitochondria were resuspended in 50 µL of *GUS* buffer (Jefferson et al., 1987), and half (~9 µg of protein) was used to assay *GUS* activity. The pellet obtained from the first low-speed spin after breaking the cells was kept to provide the chloroplast-enriched fraction. The *GUS* activity values were corrected for pigment (which absorbed some of the fluorescence) and protein content to give specific activities that could be compared between different fractions.

Yeast Constructs and Yeast Transformation

A BamHI-HindIII fragment from pYCXII (Allison and Schatz, 1986), containing the presequence-less yeast *COXIV* gene under the control of the alcohol dehydrogenase promoter, was cloned into the vector pFL38 (URA3⁺; Bonneaud et al., 1991), which had been modified previously by filling in the EcoRI site (position 396). This construct was named p Δ COX and was a gift from Françoise Budar (Institut National de la Recherche Agronomique, Versailles, France). The AlaRS presequence was amplified from the 1WT2WT construct (see above) by PCR, using the T7 primer (Stratagene) and the oligonucleotide 5'-AGC-TCTAGAGACTGCGACGGAGGA-3' bearing an XbaI site (underlined). The 250-bp PCR product was digested by Apol and XbaI restriction enzymes and inserted in p Δ COX between the EcoRI and XbaI sites, creating the p Δ COX-AlaRSp construct. The p Δ COX-Tnt-AlaRSp construct, containing the Tnt1 leader in addition, was made similarly, except that the PCR used the TntWT plasmid as substrate, and the product was digested with SspI and XbaI before ligation between the EcoRI (filled-in) and XbaI sites of p Δ COX. To make p Δ COX-Tnt-1mAlaRSp, in which the first ATG was changed to a GCA codon, a first PCR was performed on the 1WT2WT construct using the primer bearing the XbaI site (see above) and an oligonucleotide introducing the mutations (underlined) 5'-TGGTGAAAATAAAAAGCAAGATTAGTGAAGGCA-3'. One-four hundredth of the 167-bp PCR product obtained was mixed with 10 ng of pBMCV102120 plasmid for a second PCR amplification by using the same primer bearing the XbaI site (see above) and a *GUS*-specific primer 5'-GCTTCCACCAACGCTGAT-3'. The fusion product was then digested with SspI and XbaI before ligation between the EcoRI (filled-in) and XbaI sites of p Δ COX.

All constructs were used to transform the yeast strain WSR, in which the endogenous *COXIV* copy is disrupted by the *LEU2* marker. Yeast transformation was performed using the protocol described by Schiest and Giesz (1989). Tests of complementation of respiratory deficiency were made on medium lacking sugars and containing 2% glycerol.

ACKNOWLEDGMENTS

We thank Andrée Lepingle and Jan Drouaud for help with sequencing, Olivier Grandjean for the photography, and Tracy Ripmaster, Marc Boutry, and Françoise Budar for providing plasmids and yeast strains. We also thank Paul Schimmel, Marc Boutry, and André Dietrich for their helpful comments on the manuscript.

Received February 22, 1996; accepted April 25, 1996.

REFERENCES

- Allison, D.S., and Schatz, G. (1986). Artificial mitochondrial presequences. *Proc. Natl. Acad. Sci. USA* **83**, 9011–9015.
- Amaar, Y.G., and Baillie, D.L. (1993). Cloning and characterization of the *C. elegans* histidyl-tRNA synthetase gene. *Nucleic Acids Res.* **21**, 6050–6051.
- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K., eds (1990). *Current Protocols in Molecular Biology*. (New York: John Wiley and Sons).
- Belrhali, H., Yaremchuk, A., Tukalo, M., Larsen, K., Berthet-Colominas, C., Leberman, R., Beijer, B., Sproat, B., Als-Nielsen, J., Grubel, G., Legrand, J.F., Lehmann, M., and Cusack, S. (1994). Crystal structures at 2.5 angstrom resolution of seryl-tRNA synthetase complexed with two analogs of seryl adenylate. *Science* **263**, 1432–1436.
- Biou, V., Yaremchuk, A., Tukalo, M., and Cusack, S. (1994). The 2.9 angstrom crystal structure of *T. thermophilus* seryl-tRNA synthetase complexed with tRNA^{Ser}. *Science* **263**, 1404–1410.
- Bonneaud, N., Ozier-Kalogeropoulos, O., Li, G.Y., Labouesse, M., Minvielle-Sebastia, L., and Lacroute, F. (1991). A family of low and high copy replicative, integrative and single-stranded *S. cerevisiae*/*E. coli* shuttle vectors. *Yeast* **7**, 609–615.
- Bordonaro, M., Saccomanno, C.F., and Nordstrom, J.L. (1994). An improved T1/A ribonuclease protection assay. *BioTechniques* **16**, 428–430.
- Bradford, M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. *Anal. Biochem.* **72**, 248–254.
- Buechter, D.D., and Schimmel, P. (1993). Dissection of a class-II transfer RNA synthetase—Determinants for minihelix recognition are tightly associated with domain for amino acid activation. *Biochemistry* **32**, 5267–5272.
- Carneiro, V.T.C., Pelletier, G., and Small, I. (1993). Transfer RNA-mediated suppression of stop codons in protoplasts and transgenic plants. *Plant Mol. Biol.* **22**, 681–690.
- Carneiro, V.T.C., Dietrich, A., Maréchal-Drouard, L., Cosset, A., Pelletier, G., and Small, I. (1994). Characterization of some major identity elements in plant alanine and phenylalanine transfer RNAs. *Plant Mol. Biol.* **26**, 1843–1853.
- Carter, C.W. (1993). Cognition, mechanism, and evolutionary relationships in aminoacyl-tRNA synthetases. *Annu. Rev. Biochem.* **62**, 715–748.
- Cavarelli, J., Rees, B., Ruff, M., Thierry, J.C., and Moras, D. (1993). Yeast transfer RNA^{Asp} recognition by its cognate class-II aminoacyl-transfer RNA synthetase. *Nature* **362**, 181–184.
- Cavarelli, J., Eriani, G., Rees, B., Ruff, M., Boeglin, M., Mitschler, A., Martin, F., Gangloff, J., Thierry, J.C., and Moras, D. (1994). The active site of yeast aspartyl-tRNA synthetase—Structural and functional aspects of the aminoacylation reaction. *EMBO J.* **13**, 327–337.
- Cavener, D.R., and Stuart, S.C. (1991). Eukaryotic start and stop translation sites. *Nucleic Acids Res.* **19**, 3185–3192.
- Chang, P.K., and Dignam, J.D. (1990). Primary structure of alanyl-tRNA synthetase and the regulation of its mRNA levels in *Bombyx mori*. *J. Biol. Chem.* **265**, 20898–20906.
- Chatton, B., Walter, P., Ebel, J.P., Lacroute, F., and Fasiolo, F. (1988). The yeast *VAS1* gene encodes both mitochondrial and cytoplasmic valyl-tRNA synthetases. *J. Biol. Chem.* **263**, 52–57.
- Chaumont, F., O'Riordan, V., and Boutry, M. (1990). Protein transport into mitochondria is conserved between plant and yeast species. *J. Biol. Chem.* **265**, 16856–16862.
- Chaumont, F., de Castro Silva Filho, M., Thomas, D., Leterme, S., and Boutry, M. (1994). Truncated presequences of mitochondrial F₁-ATPase β subunit from *Nicotiana plumbaginifolia* transport CAT and GUS proteins into mitochondria of transgenic tobacco. *Plant Mol. Biol.* **24**, 631–641.
- Colas des Francs-Small, C., Ambard-Bretteville, F., Small, I., and Remy, R. (1993). Identification of a major soluble protein in mitochondria from non-photosynthetic tissues as NAD-dependent formate dehydrogenase. *Plant Physiol.* **102**, 1171–1177.
- Creissen, G., Reynolds, H., Xue, Y., and Mullineaux, P. (1995). Simultaneous targeting of pea glutathione reductase and of a bacterial fusion protein to chloroplast and mitochondria in transgenic tobacco. *Plant J.* **8**, 167–175.
- Cusack, S., Härtlein, M., and Leberman, R. (1991). Sequence, structure and evolutionary relationships between class 2 aminoacyl-tRNA synthetases. *Nucleic Acids Res.* **19**, 3489–3498.
- Damerval, C., de Vienne, D., Zivy, M., and Thiellement, H. (1986). Technical improvements in two-dimensional electrophoresis increases the level of genetic variation detected in wheat seedling proteins. *Electrophoresis* **7**, 52–54.
- Danpure, C.J. (1995). How can the products of a single gene be localized to more than one intracellular compartment. *Trends Cell Biol.* **5**, 230–238.
- Davis, M.W., Buechter, D.D., and Schimmel, P. (1994). Functional dissection of a predicted class-defining motif in a class II tRNA synthetase of unknown structure. *Biochemistry* **33**, 9904–9911.
- Dellaporta, S.L., Wood, J., and Hicks, J.B. (1983). A plant DNA miniprep: Version II. *Plant Mol. Biol. Rep.* **1**, 19–21.
- Dietrich, A., Weil, J.H., and Maréchal-Drouard, L. (1992). Nuclear-encoded transfer RNAs in plant mitochondria. *Annu. Rev. Cell Biol.* **8**, 115–131.
- Dumas, J.B., Edwards, M., Delort, J., and Mallet, J. (1991). Oligodeoxyribonucleotide ligation to single-stranded cDNAs: A new tool for cloning 5' ends of mRNAs and for constructing cDNA libraries by in vitro amplification. *Nucleic Acids Res.* **19**, 5227–5232.
- Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990). Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* **347**, 203–206.

- Filley, S.J., and Hill, K.A.W. (1993). Amino acid substitution at position 73 in motif 2 of *Escherichia coli* alanyl-tRNA synthetase. *Arch. Biochem. Biophys.* **307**, 46–51.
- Gabriel, K., Schneider, J., and McClain, W.H. (1996). Functional evidence for indirect recognition of G·U in tRNA^{Ala} by alanyl-tRNA synthetase. *Science* **271**, 195–197.
- Grandbastien, M.-A., Spielmann, A., and Caboche, M. (1989). Tnt1, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. *Nature* **337**, 376–380.
- Hou, Y.M., and Schimmel, P. (1988). A simple structural feature is a major determinant of the identity of a transfer RNA. *Nature* **333**, 140–145.
- Hou, Y.M., and Schimmel, P. (1989). Evidence that a major determinant for the identity of a transfer RNA is conserved in evolution. *Biochemistry* **28**, 6800–6804.
- Jasin, M., Regan, L., and Schimmel, P. (1983). Modular arrangement of functional domains along the sequence of an aminoacyl-tRNA synthetase. *Nature* **306**, 441–447.
- Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W. (1987). GUS fusions: β -Glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6**, 3901–3907.
- Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283–292.
- Logemann, J., Schell, J., and Willmitzer, L. (1987). Improved method for the isolation of RNA from plant tissues. *Anal. Biochem.* **163**, 16–20.
- Lu, Y., and Hill, K.A.W. (1994). The invariant arginine in motif 2 of *Escherichia coli* alanyl-tRNA synthetase is important for catalysis but not for substrate binding. *J. Biol. Chem.* **269**, 12137–12141.
- Maréchal-Drouard, L., Guillemaut, P., Cosset, A., Arbogast, M., Weber, F., Weil, J.H., and Dietrich, A. (1990). Transfer RNAs of potato (*Solanum tuberosum*) mitochondria have different genetic origins. *Nucleic Acids Res.* **18**, 3689–3696.
- Martin, N.C., and Hopper, A.K. (1994). How single genes provide tRNA processing enzymes to mitochondria, nuclei and the cytosol. *Biochimie* **76**, 1161–1167.
- McClain, W.H. (1993). Rules that govern tRNA identity in protein synthesis. *J. Mol. Biol.* **234**, 257–280.
- Moras, D. (1992). Structural and functional relationships between aminoacyl-tRNA synthetases. *Trends Biochem. Sci.* **17**, 159–164.
- Mukumoto, F., Hirose, S., Imaseki, H., and Yamazaki, K. (1993). DNA sequence requirement of a TATA element-binding protein from *Arabidopsis* for transcription *in vitro*. *Plant Mol. Biol.* **23**, 995–1003.
- Musier-Forsyth, K., Usman, N., Scaringe, S., Doudna, J., Green, R., and Schimmel, P. (1991). Specificity for aminoacylation of an RNA helix—An unpaired, exocyclic amino group in the minor groove. *Science* **253**, 784–786.
- Natsoulis, G., Hilger, F., and Fink, G.R. (1986). The *HTS1* gene encodes both the cytoplasmic and mitochondrial histidine-tRNA synthetases of *S. cerevisiae*. *Cell* **46**, 235–243.
- Pouteau, S., Huttner, E., Grandbastien, M.A., and Caboche, M. (1991). Specific expression of the tobacco Tnt1 retrotransposon in protoplasts. *EMBO J.* **10**, 1911–1918.
- Putney, S.D., Royal, N.J., Neuman de Vegar, H., Herlihy, W.C., Biemann, K., and Schimmel, P. (1981). Primary structure of a large aminoacyl-tRNA synthetase. *Science* **213**, 1497–1500.
- Ribas de Pouplana, L., Buechter, D.D., Davis, M.W., and Schimmel, P. (1993). Idiographic representation of conserved domain of a class II tRNA synthetase of unknown structure. *Protein Sci.* **2**, 2259–2262.
- Ripmaster, T.L., Shiba, K., and Schimmel, P. (1995). Wide cross-species aminoacyl-tRNA synthetase replacement in vivo: Yeast cytoplasmic alanine enzyme replaced by human polymyositis serum antigen. *Proc. Natl. Acad. Sci. USA* **92**, 4932–4936.
- Saks, M.E., Sampson, J.R., and Abelson, J.N. (1994). The transfer RNA identity problem: A search for rules. *Science* **263**, 191–197.
- Schiest, R., and Giesz, R.D. (1989). High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr. Genet.* **16**, 339–346.
- Schmitz, U.K., and Lonsdale, D.M. (1989). A yeast mitochondrial presequence functions as a signal for targeting to plant mitochondria in vivo. *Plant Cell* **1**, 783–791.
- Shi, J.P., Musier-Forsyth, K., and Schimmel, P. (1994). Region of a conserved sequence motif in a class II tRNA synthetase needed for transfer of an activated amino acid to an RNA substrate. *Biochemistry* **33**, 5312–5318.
- Shiba, K., Ripmaster, T., Suzuki, N., Nichols, R., Plotz, P., Noda, T., and Schimmel, P. (1995). Human alanyl-tRNA synthetase: Conservation in evolution of catalytic core and microhelix recognition. *Biochemistry* **34**, 10340–10349.
- Slusher, L.B., Gillman, E.C., Martin, N.C., and Hopper, A.K. (1991). mRNA leader length and initiation codon context determine alternative AUG selection for the yeast gene *MOD5*. *Proc. Natl. Acad. Sci. USA* **88**, 9789–9793.
- Small, I., Maréchal-Drouard, L., Masson, J., Pelletier, G., Cosset, A., Weil, J.H., and Dietrich, A. (1992). *In vivo* import of a normal or mutagenized heterologous transfer RNA into the mitochondria of transgenic plants: Toward novel ways of influencing mitochondrial gene expression? *EMBO J.* **11**, 1291–1296.
- Stein, I., Peleg, Y., Even-Ram, S., and Pines, O. (1994). The single translation product of the *FUM1* gene (fumarase) is processed in mitochondria before being distributed between the cytosol and mitochondria in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **14**, 4770–4778.
- Steinmetz, A., and Weil, J.H. (1986). Isolation and characterization of chloroplast and cytoplasmic transfer RNAs. *Methods Enzymol.* **118**, 212–231.
- Suzuki, T., Yoshida, T., and Tuboi, S. (1992). Evidence that rat liver mitochondrial and cytosolic fumarases are synthesized from one species of mRNA by alternative translation initiation at two in-phase AUG codons. *Eur. J. Biochem.* **207**, 767–772.
- Tarassov, I.A., and Entelis, N.S. (1992). Mitochondrially-imported cytoplasmic transfer RNA^{Lys}(CUU) of *Saccharomyces cerevisiae*—In vivo and in vitro targeting systems. *Nucleic Acids Res.* **20**, 1277–1281.
- Tarassov, I., Entelis, N., and Martin, R.P. (1995a). An intact protein translocating machinery is required for mitochondrial import of a yeast cytoplasmic tRNA. *J. Mol. Biol.* **245**, 315–323.
- Tarassov, I., Entelis, N., and Martin, R.P. (1995b). Mitochondrial import of a cytoplasmic lysine-tRNA in yeast is mediated by cooperation of cytoplasmic and mitochondrial lysyl-tRNA synthetases. *EMBO J.* **14**, 3461–3471.
- von Heijne, G., Steppuhn, J., and Herrmann, G. (1989). Domain structure of mitochondrial and chloroplast targeting peptides. *Eur. J. Biochem.* **180**, 535–545.
- Wolfe, C.L., Lou, Y.C., Hopper, A.K., and Martin, N.C. (1994). Interplay of heterogeneous transcriptional start sites and translational selection of AUGs dictate the production of mitochondrial and cytosolic nuclear tRNA nucleotidyltransferase from the same gene in yeast. *J. Biol. Chem.* **269**, 13361–13366.