

Mapping of mitochondrial mRNA termini in *Arabidopsis thaliana*: t-elements contribute to 5' and 3' end formation

Joachim Forner, Bärbel Weber, Sabine Thuss, Steffen Wildum and Stefan Binder*

Institut Molekulare Botanik, Universität Ulm, Albert-Einstein-Allee 11, 89069 Ulm

Received March 7, 2007; Revised April 5, 2007; Accepted April 9, 2007

ABSTRACT

With CR–RT–PCR as primary approach we mapped the 5' and 3' transcript ends of all mitochondrial protein-coding genes in *Arabidopsis thaliana*. Almost all transcripts analyzed have single major 3' termini, while multiple 5' ends were found for several genes. Some of the identified 5' ends map within promoter motifs suggesting these ends to be derived from transcription initiation while the majority of the 5' termini seems to be generated post-transcriptionally. Assignment of the extremities of 5' leader RNAs revealed clear evidence for an endonucleolytic generation of the major *cox1* and *atp9* 5' mRNA ends. tRNA-like structures, so-called t-elements, are associated either with 5' or with 3' termini of several mRNAs. These secondary structures most likely act as *cis*-signals for endonucleolytic cleavages by RNase Z and/or RNase P. Since no conserved sequence motif is evident at post-transcriptionally derived ends, we suggest t-elements, stem-loops and probably complex higher order structures as *cis*-elements for processing. This analysis provides novel insights into 5' and 3' end formation of mRNAs. In addition, the complete transcript map is a substantial and important basis for future studies of gene expression in mitochondria of higher plants.

INTRODUCTION

Seed plants contain the largest mitochondrial genomes investigated so far. Although their sizes reach up to about half of the *Escherichia coli* genome, the mitochondrial genomes more or less encode only the same small set of ~60 genes. This gene collection is found with minor variations in the mitochondrial genome sequences of *Arabidopsis thaliana*, sugar beet, rape seed, tobacco, wheat, rice and maize (1–8). The transcription units,

mono- as well as polycistronic, are spaced across the entire genome lengths and are separated by large spacer sequences without any obvious function. Transcription of mitochondrial genes frequently starts from multiple promoters of various types, generating precursor RNAs that have to pass through various processing steps such as RNA editing, splicing of group II introns, 3' end trimming as well as formation of secondary 5' termini (9). Many or all of these steps are required to generate a mature translatable mRNA or functional transfer as well as ribosomal RNAs and can potentially contribute to the regulation of mitochondrial gene expression. But up to date it is still unknown whether or to which extent these processes control or even regulate the realization of the mitochondrial genetic information.

Transcription initiation is one of the most important levels to regulate gene expression in bacteria, archaea as well as nuclei from eukaryotes. This process has also been intensively examined in mitochondria of various plant species. In *A. thaliana* for instance at least two conserved promoter motifs have been found and in addition a number of non-conserved transcription initiation sites are present (10–12). The situation might be even more complex since not all mitochondrial transcription units have been analyzed in this respect. Up to date it is still unclear to what extent if at all plant mitochondrial gene expression is regulated or controlled at the transcriptional level. Run on transcription studies showed that mitochondrial genes are transcribed at different rates, most likely determined by differing promoter strength (13,14). But so far there is no clear evidence that expression of individual genes is actively regulated during transcription initiation events.

Several reports provided convincing evidence that also post-transcriptional processes influence plant mitochondrial mRNA steady state levels (15,16). This was demonstrated by comparing transcriptional rates with the steady state RNA levels. The observed discrepancies for several genes were interpreted to originate from post-transcriptional processes influencing RNA stability. Still little is known how the stability of a transcript is determined.

*To whom correspondence should be addressed. Tel: +49 731 5022625; Fax: +49 731 5022626; Email: stefan.binder@uni-ulm.de

Of course there must be *cis*-elements as well as *trans*-factors, which probably determine RNA stability in a concerted action. Stem-loop (SL) structures are good candidates for *cis*-acting processing signals and such structures at or near the 3' ends of several plant mitochondrial RNAs have been found to influence the mRNA stability *in vivo* and *in vitro* (17–20), most likely preventing exonucleolytic degradation (21). However, the nature or features of structures determining 3' ends and/or RNA stability are still unknown. Another important (*cis*)-factor is the polyadenylation state on an RNA. Short oligo(A) tails at mature 3' ends have been found to destabilize plant mitochondrial RNA both *in vivo* and *in vitro* and are thus expected only at a minor fraction of the steady state pool (9,18,22,23).

The 5' ends of mRNAs can be generated directly by transcription initiation or by subsequent 5' processing events. Support for the existence of the latter has been obtained by mapping of such ends of various genes in mitochondria of different plant species. It is still unclear however, how these ends are generated. Up to now no evidence has been reported for a 5' to 3' exonucleolytic activity. Consequently the generation of secondary 5' termini has been attributed to (an) endoribonuclease(s).

So far two different endonucleolytic RNase activities have been described in plant mitochondria, both being involved in the maturation of tRNAs. An RNase P-like activity has been found to cut precisely at the mature tRNA 5' end, while RNase Z cleaves directly at or one nucleotide downstream of the discriminator nucleotide at the 3' end (24–26). The prerequisite for the cleavage of precursor molecules by these activities is the formation of the tRNA secondary structure (27,28). tRNA-like elements (t-elements) with non-canonical cloverleaf structures are also substrates for these enzymes at least *in vitro*. Such t-elements have been found dispersed in wheat mtDNA and if transcribed as part of an RNA molecule these structures could potentially be involved in cleavage of such long primary transcripts to generate secondary ends (29,30). But so far only little evidence has been found for the function of such t-elements *in vivo*.

To gain more information about 5' and 3' ends of mRNAs in plant mitochondria, we analyzed mRNA extremities of all protein-coding genes annotated in the mitochondrial genome of *A. thaliana*. Single 3' ends are found for almost all transcription units, while multiple 5' ends are found in several cases. Only a few of the 5' ends are found at conserved promoter sequences, while most 5' termini are most likely generated post-transcriptionally. Analysis of the sequences surrounding the 5' and 3' termini revealed that several of the ends coincide with termini of t-elements or stem-loop structures suggesting that RNase Z and RNase P are involved in mRNA 5' and 3' end processing.

MATERIALS AND METHODS

Preparation of RNA from *A. thaliana*

An *A. thaliana* ecotype Columbia cell suspension culture was cultivated on a shaker (120 r.p.m.) in the dark at 23°C.

The ecotype of this culture was recently confirmed by analyzing corresponding informative genetic markers (31). Mitochondria were isolated from cultures six days after the transfer to fresh medium according to a method described previously (32). RNA from these organelles (mtRNA) was extracted following previously established protocols (33). Alternatively, 100 mg frozen mitochondria (fresh weight) were disrupted in a mortar. The fractured organelles were suspended in the lysis buffer of an RNeasy Plant Mini kit. RNA was isolated following the instructions of the manufacturer (Qiagen).

Total cellular RNA was isolated from 100 mg cell suspension culture. To this end cells were harvested 24 h after transfer to fresh medium and were ground in liquid nitrogen. From this material total RNA was purified using an RNeasy Plant Mini kit as above.

Analysis of RNA

CR-RT-PCR analyses were either performed as described before (Figure 1A) (34) and/or carried out using a modified protocol (Figure 1B). Briefly, large-scale self-ligation was performed with up to 50 µg RNA in a total volume of 100 µl. After ligation, samples were desalted using Microcon YM-10 or -30 micro concentrators (Millipore) and stored as aliquots of 15 µl at -20°C. First strand cDNA synthesis was done with 5 µg of total RNA and 2 µg of mitochondrial RNA, respectively, and 200 U M-MLV RNase H Minus (point mutant) reverse transcriptase under conditions recommended by the manufacturer (Promega). The RNA template was then degraded by adding 1/5 volume of 1 M NaOH and an incubation of 10 min at room temperature. The sample was subsequently neutralized with an equal amount of 1 M HCl and the cDNA was purified with the GFX™ PCR DNA and Gel Band Purification Kit (GE Healthcare). This purification step also removed the primer used for cDNA synthesis. This oligonucleotide could cause the amplification of multiple PCR products when present in the first amplification reaction containing another primer with the same orientation. One-fifth of the cDNA sample was used as template in a single RT-PCR. In some cases amplification was done in a two-step PCR using primers with melting temperatures of about 73°C. This allows annealing and synthesis to be performed at 68°C.

The extremities of the *nad6*, *ccmC* and *cox1* t-elements, respectively, were determined by using mitochondrial 5S and 18S rRNAs as linker molecules, since 5' and 3' end mapping of small RNA species with conventional CR-RT-PCR is rather inefficient. The rRNAs were linked to the target RNAs by bulk ligation of 5 µg of mitochondrial RNA as described above. The use of 18S instead of 5S rRNA as anchor molecule permits an amplification of larger products, which increases PCR efficiency. Further details are given in Figure 6 and Supplementary Figures 29, 30 and 31.

For northern blot experiments 3–10 µg of total or 1–5 µg of mitochondrial RNA were size fractionated on 1% (w/v) agarose gels using glyoxal as denaturing agent. Gel electrophoresis, blotting and hybridization were

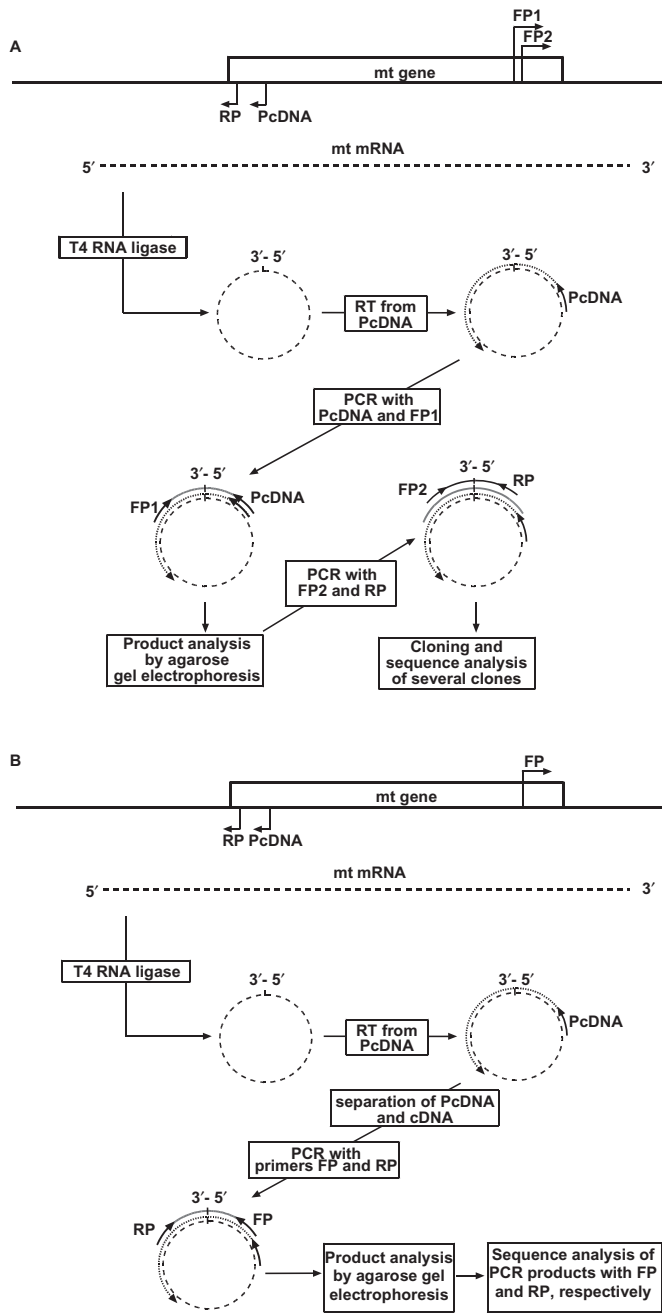


Figure 1. CR-RT-PCR analysis of mitochondrial mRNAs. Two different CR-RT-PCR strategies were followed in the transcript analysis of the 32 protein-coding genes encoded in the mitochondrial genome of *A. thaliana*. The main difference is that the approach outlined in (A) includes two PCRs and cloning with subsequent sequence analysis of individual clones while in the alternative approach (B) PCR products obtained by a single PCR step were directly sequenced. Further explanations see in the text. FP: Forward primer; RP: reverse primer; PcDNA: oligonucleotide used to prime cDNA synthesis.

done as described previously (35). Nucleic acids were blotted onto Duralon UV (Stratagene) or Hybond-XL membranes (GE Healthcare) and hybridized with radioactively labeled probes as outlined in the manufacturers' guidelines. Primer extension analyses were performed according to standard protocols (35).

Miscellaneous methods

Sequencing of PCR products was commercially obtained (MWG Biotech and 4baselab). Individual cDNA clones were sequenced using Thermo Sequenase Primer Sequencing kits as recommended by the manufacturer (GE Healthcare) and an ALF sequencing system (GE Healthcare). *In silico* sequence analyses were done at the NCBI server using various blast tools (36). Secondary structure predictions were done with the program 'stemloops' from the Wisconsin GCG software package version 10.2, with a program searching for palindromes (<http://bioweb.pasteur.fr/seqanal/interfaces/palindrome.html>) and a web interface RNA-fold program (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>). Searches for conserved sequence motifs were done with the MEME search tools (<http://meme.nbcr.net/meme/meme.html>).

RESULTS

The experimental approach to map major mRNA termini in *A. thaliana* mitochondria

The *A. thaliana* mitochondrial genome encodes 32 protein-coding genes. To gain more information about the major termini of the transcripts of these genes in this model plant species we analyzed the RNAs covering these loci by CR-RT-PCR. This experimental approach allows the simultaneous determination of 5' and 3' termini, the detection of non-encoded nucleotides and delivers unambiguous information by sequencing of the amplified cDNA products. In addition this method allows fine mapping of the ends on the nucleotide level.

Generally two different approaches were used, both starting from gene-specific cDNA synthesized on self-ligated RNA (Figure 1). In the first approach, which was mainly used in the initial phase of the project, products were amplified from mitochondrial RNA isolated from *A. thaliana* ecotype Columbia (Col) cell suspension culture. A first PCR was performed with the primer used for cDNA synthesis, which anneals in the 5' terminal region of the reading frame and a forward primer complementary to sequences located in the 3' terminal part of the gene. Products of this reaction were inspected on agarose gels and the prominent fragments were selected for further amplification in (a) nested PCR(s). The resulting fragment(s) was (were) cloned and about 20 cDNA clones were sequenced for each PCR product (Figure 1A).

In the alternative approach, gene-specific cDNA synthesis was performed on both mitochondrial and total RNA from the Col cell suspension culture. After cDNA synthesis a single PCR was performed with oligonucleotides annealing to 5' and 3' terminal sequences of the reading frame (Figure 1B). In this PCR, the primer annealing to the 5' part of the gene is different from the oligonucleotide used for cDNA synthesis. In some cases primers with melting temperatures above 73°C were used to increase the specificity in these reactions. Prominent products which appeared identical in the PCR patterns of both RNAs were selected for sequencing from both

cDNA synthesis was initiated from oligonucleotide Atcox1-1 complementary to sequences from position +156 to +138 relative to the ATG (+1). The same primer and oligonucleotide Atcox1-3 (+1473 to +1493) were used in the first amplification reaction yielding a strong product of 500 and two weak products with sizes of about 400 and 650 nucleotides, respectively (Figure 2A and B, left panel). These were subsequently used as DNA templates in three separate amplifications (second PCR), each performed with primer pair Atcox1-2 (+118 to +98)/Atcox1-4 (+1501 to +1520). While no product was obtained in the reaction with the smallest DNA template, strong fragments of 500 and about 300 nucleotides were generated in the reaction with the 500-bp template. A relatively weak cDNA product of 500 bp is obtained from the 650-bp template. The 500-bp cDNA from the second PCR was then cloned, which yielded cDNA clones with sizes between 270 and 500 bp (data not shown). Sequence analysis of 32 clones of various sizes identified 3' ends in 23 clones 47, in six clones 46 and in one clone 45 nucleotides downstream of the *cox1* stop codon, respectively (Supplementary Table 1). A single cDNA clone has a 3' end in the reading frame 40 nucleotides upstream of the stop codon while another clone has a 3' end 64 nucleotides downstream of the *cox1* reading frame. This analysis suggests a single dominant 3' end for all *cox1* mRNAs 47 bp downstream of the reading frame with some slightly scattering 3' ends around this position (349 783 in complete sequence nc_001284.2). Non-encoded nucleotides, i.e. four adenines followed by a fragment of the mitochondrial 18S rRNA were found in a single cDNA clone.

In contrast to the homogenous 3' terminus, the 5' ends found in individual cDNA clones vary considerably. The 5' termini in the largest cDNA inserts are found in five clones 241, in two clones 240 and in three clones 239 nucleotides upstream of the *cox1* ATG start codon. All other 20 clones have 5' ends between -216 to -20 relative to the ATG with two minor clusters between -187 and -179 and -48 and -43. This suggests that a major 5' terminus of the steady state *cox1* mRNA is located 241 nucleotides upstream of the ATG (at position 351 654). The other clones are either minor ends or represent degradation products.

In the second approach, cDNA synthesis was again initiated from primer Atcox1-1 but was followed by a single amplification reaction with primer pair Atcox1-2/Atcox1-4. A 500-bp cDNA was obtained as dominant product from both RNAs (Figure 2C). This PCR product was directly sequenced with oligonucleotide Atcox1-2 directed to the 5' end, yielding a sequence across the ligation sites of the *cox1* steady state mRNA pool (Figure 2D). The chromatogram shows a clean sequence up to the adenosine at position 351 654 followed by an overlay of minor signals at all positions but with a dominant unique sequence starting with an adenosine at position 349 783. This transition marks the ligation site and determines the 5' and 3' transcript termini. The majority of the *cox1* transcripts thus starts with a thymidine 241 nucleotides upstream of the ATG and ends with another thymidine 47 nucleotides downstream

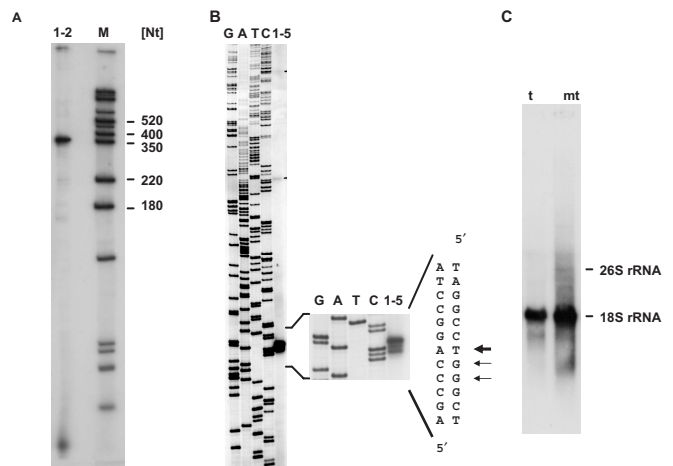


Figure 3. Primer extension and northern blot analysis of *cox1* transcripts. (A) Extension products generated in a reverse transcription reaction from oligonucleotide Atcox1-2 (Figure 2A) are separated along DNA marker fragments. A major product (black arrow) detects a 5' end about 240 nucleotides upstream of the ATG, which is consistent with the major 5' terminus mapped in the CR-RT-PCR analysis. (B) The major end(s) is (are) also detected in the extension reaction with oligonucleotide Atcox1-5. The corresponding products (lanes 1-5) were separated together with sequencing reaction products (lanes G, A, T and C) obtained with the same primer. These ends (black arrows) scatter over three nucleotides with the most prominent end found at -241 and minor termini at -240 and -239. The sequence is given in both orientations. (C) Northern blot analysis performed with a probe covering most of the *cox1* reading frame as indicated in Figure 2A. A single mRNA of about 1900 nucleotides is detected in total (t) and mitochondrial (mt) RNA preparations from *A. thaliana* cell suspension culture of ecotype Col.

of the reading frame. That the sequence can be clearly followed across the ligation site indicates that the vast majority of the *cox1* mRNA molecules has identical 5' and 3' ends. The appearance of minor signals beyond the ligation site shows the presence of a minor fraction of *cox1* RNA molecules with slightly differing 5' ends. Both approaches consistently identify the dominant *cox1* mRNA 5' end located 241 nucleotides upstream of the ATG and a 3' terminus corresponding to a 47 nucleotide 3' UTR.

To check these CR-RT-PCR mapping results by independent methods, primer extension reactions were carried out with primers Atcox1-2 and Atcox1-5 (-129 to -148) using mtRNA as template. Atcox1-2 is elongated to a major product of about 360 nucleotides corresponding to the above detected 5' end at -241. Additional weak products indicate minor ends up- and downstream of the main terminus (Figure 3A). Higher resolution by separation of Atcox1-5 extension products alongside sequencing reaction products from the same primer shows major products at positions -241 to -239 with the same quantitative distribution as seen in the CR-RT-PCR (Figure 3B). For further analysis by yet another independent method a northern blot experiment was performed. This assay detects a single transcript of about 1900 nucleotides in each RNA preparation consistent with the size of 1872 nucleotides calculated on the basis of the mapping data (Figure 3C).

In summary, these results clearly show that in *A. thaliana* the major *cox1* transcripts range from 241 bp upstream of the ATG to 47 bp downstream of the stop codon consistent with an mRNA of 1872 nucleotides. Thus, the major 5' terminus is located 114 bp downstream of the previously identified promoter 355 bp upstream of the ATG (12). In addition, these data document again the reliability of the CR-RT-PCR analysis, particularly of the approach with direct sequencing of PCR products.

Single 3' ends are detected for almost all *A. thaliana* mitochondrial mRNAs

Using the experimental approaches shown in Figure 1 we analyzed the transcripts of the mitochondrially encoded proteins in *A. thaliana*. Details of the individual mapping procedures and their results are shown in the gene-specific supplementary figures (Supplementary Figures 1–27). These indicate the primers used for cDNA synthesis as well as the amplification reactions and document the results of the PCRs including the positions of the 5' and 3' ends. The two CR-RT-PCR strategies detect those ends, which are most abundant and which are located within distances of a few hundred nucleotides from the primers used. This assumption was independently confirmed for many transcripts, either by northern or primer extension analyses for instance for *cox1* and *atp9*.

All 3' termini detected are given in Table 1. For all mRNAs single slightly scattering 3' ends are found, the only exception being the *ccmF_C* transcript for which two ends were detected. One of these *ccmF_C* ends (around +115) shows an unusually wide scattering over several nucleotides (Table 1).

The identified 3' ends are mostly located downstream of the reading frame with 3' UTRs up to 498 nucleotides found for *rps4*. The major 3' termini of the *ccmC* and *nad6* genes are found within these reading frames (*ccmC*: –46 and *nad6*: –17 before the stop codon) confirming recently published results (37). Additional 3' ends downstream of these reading frame-internal major 3' ends can be detected only at particular processing products.

Two other 3' termini were identified within the two pseudo genes ψ *mttB* and ψ *sdh4*, the latter corresponding to the 3' end of the *cox3* mRNA (38).

A MEME analysis of sequences flanking the 3' ends (–/+ 20 bp) did not reveal any conserved sequence motif. However, it is noteworthy that transcripts ending with a guanidine are under-represented. Nine transcripts each end with cytidines and uridines, respectively, while the mRNAs of six genes end with adenosines, which is roughly the random distribution.

Sequences flanking the 3' termini were also screened for the presence of inverted repeats with the potential to form stem-loop structures (SL). The most striking structure is a double stem-loop found exactly upstream of the 3' end of the *atp9* mRNA (Figure 4). This has been observed in a previous study, which suggested this structure to have a function in the exonucleolytic maturation of this end (21). Parts of this double-inverted repeat are duplicated in the 3' flanking region of *nad1* exon e with the downstream-located SL being identical. The mature *nad1* transcript

termini being at identical positions as the *atp9* mRNA 3' ends indicate a common function of this duplicated SL in formation of these 3' ends. A single stem-loop with different primary sequence is found in an analogous position directly upstream of the *cox2* 3' terminus, which suggests a similar function of this SL in the generation of this end (Figure 4). Inverted repeats in the 3' UTRs of the *cox1*, *nad2* and *rps7* genes form single SLs and are positioned several nucleotides upstream of these ends (data not shown). Two SLs separated only by 10 nucleotides are present in the *atp8* 3' UTR. Interestingly the 3' terminus of the latter transcript maps one nucleotide downstream of the 3' end of a short stable RNA (Ath-377) identified in a general RNomics analysis in *A. thaliana* (39). Similarly the 3' end of the *nad4L-atp4* mRNA coincides with the terminus of Ath-290.

5' ends of mitochondrial transcripts in *A. thaliana*

The 5' termini identified in our analysis are listed in Table 2. These ends could originate directly from transcription initiation (referred to as primary ends) or could be formed by post-transcriptional processing (known as secondary ends). The mechanism of RNA ligation allows only the connection of 5' monophosphate ends and excludes 5' termini with two or three phosphates which are expected at primary ends derived from transcription initiation (40,41). However, the triphosphate ends are rather unstable and thus primary ends can also carry 5' monophosphate groups, which allows their direct ligation and detection without treatment of the RNA. This was likewise observed previously (12).

Several major 5' ends detected in our experiments have been identified previously as primary ends in *A. thaliana* mitochondria (12). In agreement with the results of this study we identified identical 5' termini for *atp6-1* (–200), *atp8* (–157), *atp8* (–228) and *cox2* (–205) mRNAs. For *atp8* we found an additional minor end at position –224 instead of –226 as detected before. The major 5' end of the *ccmB* mRNA at position –140 could as well be of primary origin, since it maps within a perfectly conserved promoter motif (12). In addition, a minor end of *atp1* (–1898) was found at the predicted site within the CNM1 promoter motif. This terminus was only observed in a CR-RT-PCR experiment designed to identify mRNA ends in this specific region (Supplementary Figure 1D and E).

Considering the variety of different promoter sequences found in *A. thaliana* mitochondria (10–12) a clear assignment of some of the ends identified in this study is difficult. Nevertheless, it is reasonable to assume that most of the 5' ends detected are the result of post-transcriptional processes since the surrounding sequences do not show any similarity to previously characterized promoter motifs (10,12,42–45).

Of the 30 major ends potentially derived from post-transcriptional processing only four transcripts start with a cytidine, seven with guanidine, nine with uridine and 10 with adenosine. As for the 3' ends no conserved general sequence motifs emerge at or within the sequence flanking the secondary ends, while the MEME analysis

Table 1. 3' mRNA termini in *A. thaliana* mitochondria

Gene	Position relative to the stop codon ^a (position in nc_001284)	Flanking sequences (5' to 3'), the most prominent ends are underlined with the most dominant end given in a large letter	Remarks
<i>atp6-1</i>	+45 (112 952)	TTATAATTGAACAAAAGCGA <u>G</u> GGATGGATGCTCGAGCGGTT	The large G corresponds to the position one nucleotide upstream of the tRNA ^{Ser} 5' end, 3' mRNA end generated by RNase P
<i>atp6-2</i>	+45 (296 775)	TTATAATTGAACAAAAGCGA <u>G</u> GGATGGATGCTCGAGCGGTT	The large G corresponds to the position one nucleotide upstream of the tRNA ^{Ser} 5' end, 3' mRNA end generated by RNase P
<i>ccmC (ccb3)</i>	-46 (240 033)	ATACCAAAACCTAGTTCAC <u>T</u> CGCTCTCTTTGCATCCATGGC	Major 3' end within the reading frame (37), most likely generated by the participation of RNase P
<i>nad6</i>	-17 (76 658)	AATGAGGAGGACGACTGACC <u>C</u> ACTCAGCATCTACTAAAGGG	Major 3' end within the reading frame (37), most likely generated by the participation of RNase P
<i>nad4L-atp4 (orf25)</i>	+72 (188 012)	GGATCCAATACCAAGACTAC <u>T</u> TCTTTCTCAGGAAGTGCAGC	3' end corresponds to the 3' terminus of Ath-290 (39)
<i>atp8</i>	+121 (130 506)	CTTTCTTTTCAGGCTTGACT <u>C</u> ATTTCGCTAGCTTCTCCTCCC	3' end one nucleotide downstream of Ath-377 (39)
<i>nad1</i>	+88 (143 131)	TTAATCAAGCAAGTTGGGGA <u>A</u> CAAAATCTTCTTGTTAGTT	Identical with the 3' terminus of the <i>atp9</i> mRNA
<i>atp9</i>	+180 (279 332)	TAATCAAGCAAGTTGGGGA <u>A</u> CAAAATCTTCTTGTTAGTTA	Identical with previously identified 3' terminus (21)
<i>rpl2-mitB (orfX)</i>	-10 (158 342)	CCTAGAACCTGGCAAAGTAA <u>C</u> TATCAATGAATTCCAATAAA	
<i>nad3-rps12</i>	+15 (260 209)	ATATGAATGGAAGATGCCTC <u>T</u> GGAACTAGTATCGGTCAGTC	
<i>nad4</i>	+30 (169 704)	GCAAAGAAAAGAAAACGGG <u>T</u> CAACATCTTAATGTGTATTT	
<i>nad2</i>	+35 (327 855)	AAGGTTCAAAGAAAAGGGTAG <u>G</u> CCGTCGGTTTGAATCCAATA	
<i>cox1</i>	+47 (349 783)	TAAGAACCTAACAGAACTTT <u>T</u> AGAAATTTATGCATGGTTAT	
<i>nad7</i>	+52 (138 205)	TTATTGCGAGCCCAGAAGTC <u>T</u> CTCTTTTTTTCGGCCTTCAG	
<i>nad9</i>	+55 (23 608)	CAAATCAATAGGAAATGCTA <u>T</u> AGGAAATGCTATTTGCTTTG	
<i>rpl5-cob</i>	+58 (61 474)	CACCAAGAATTGACAAGCAG <u>A</u> TAAAGTTTCTAGTTTGCTAT	
<i>rps3-rpl16</i>	+67 (25 009)	AAGTAAGGACCAAGGATAAT <u>C</u> TTTTCGGAAAGGAGGAGTAG	
<i>ccmB (ccb2)</i>	+81 (31 164)	ATGTAAGCCATGTATCTAGG <u>A</u> GGAAATTGAAAGAAGGGCTT	
<i>nad5</i>	+86 (20 485)	TCTTTCTCTCTCTCTTTTT <u>T</u> TAGTTTGTGACAGCTCATC	
<i>ccmF_C (ccb6c)</i>	+115 (51 178)	AATAGAAAATCGTATTTCGTGA <u>A</u> TAAATCCCTTTGTTTGAAT	
<i>atp1</i>	+160 (302 006)	TGGCATGCTCCCCAGGGCC <u>C</u> TCTTTTTTTAGATAGGTAGG	
<i>ccmF_{N2} (ccb6n2)</i>	+170 (256 695)	TTCGGAAGATGGTCAAGGTA <u>C</u> GAAAGTACAAGCCACTGCGC	
<i>cox3</i>	+314 (219 391)	TTGAAGAACCCTAATGGATCGA <u>A</u> ACTCATCCTACGAAAGATAT	
<i>cox2</i>	+316 (40 186)	CGGCCTACTTATAGCCTTTT <u>C</u> CCCTTTATTCGGGGACTTCA	
<i>rps7</i>	+351 (315 424)	CACTTCAAAGGGAGGGAAGG <u>C</u> TAGGAATCTCGTTTTTGCTC	
<i>ccmF_{N1} (ccb6n1)</i>	+468 (231 426)	GGTCCAGTTCTTTGCTAGTA <u>A</u> TGAGGAAAAATATATCACAC	
<i>rps4</i>	+498 (81 530)	CAAAAGGACCAAGGACGATC <u>T</u> TTTTCGGAAAGGAGGAGTAAG	
<i>matR</i>		No defined 3' end detectable	
<i>ccmF_C (ccb6c)</i>	+10 (51 283)	TAATTGCATAAAAGGAGTCA <u>A</u> AGTAGTGGCTGCGGCGCGTC	minor 3' end

^aThe first nucleotide downstream of the stop codon corresponds to position +1.

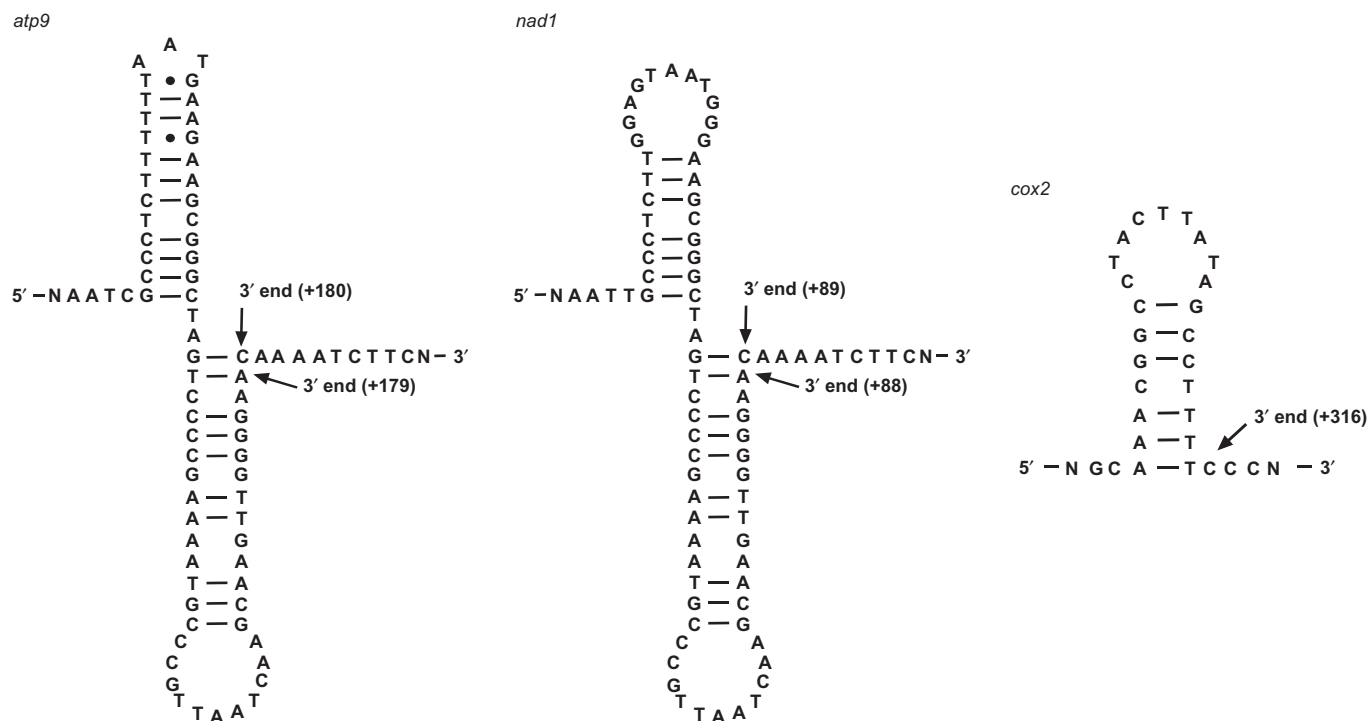


Figure 4. Stem-loop structures found directly upstream of the mapped 3' ends. The termini are indicated by arrows.

identified the promoter motif when the primary ends were analyzed (data not shown). However, we notice a striking similarity between the primary structures surrounding the 5' terminal nucleotides of the 26S rRNA, the *nad6* and the *atp9* mRNAs (Figure 5). Here 24 nucleotides are perfectly conserved between 26S rRNA and *nad6*, of which 22 are identical also in the *atp9* sequence. This suggests that these primary sequences might be important for the generation of these ends and/or that they are part of a conserved secondary structure element involved in signaling this processing.

The 5' UTRs extend up to 645 nucleotides (*nad1*) as described earlier. The mRNA of *rps4* has no 5' UTR at all, the major 5' terminus being identified at position +2 within the ATG. Since an approach to map further upstream-located 5' ends failed, it remains unclear whether this gene is functional or not.

Potential dicistronic transcripts

The *A. thaliana* mitochondrial genome encodes a number of genes that are potentially transcribed within dicistronic transcripts. This includes *rpl5-cob*, *rpl2-mttB*, *nad3-rps12*, *nad4L-atp4* and *rps3-rpl16*. All these genes were examined by CR-RT-PCR specifically assaying for monocistronic mRNA or dicistronic transcripts. Clear PCR products and ends were found for potential monocistronic mRNAs of *rpl2* (3' end at +106), while no distinct ends were detected for monocistronic transcripts of the other nine genes found within tandem arrangements. The CR-RT-PCR assays of the downstream genes of the *nad3-rps12* and *nad4L-atp4* arrangements detected 5' ends exclusively upstream of the 5' located genes suggesting a

co-transcription of these genes. Likewise, the analysis of the upstream gene of *rps3-rpl16* detected a 3' end exclusively downstream of the 3' located gene.

The potentially dicistronic genes were also investigated by CR-RT-PCR assays in which one primer annealed to the 5' terminal region of the upstream reading frame and the other to the 3' terminal part of the downstream-located gene. In all instances, these primer combinations yielded distinct and clean PCR product. This together with the lack of clear ends of potential monocistronic RNAs suggests dicistronic transcription for these genes. In the case of *rpl2* additional monocistronic mRNAs seem to be present in the steady state RNA pool, while no such transcript is detectable for *mttB*.

Endonucleolytic cleavage generates the 5' end of the mature *cox1* mRNA

As indicated in the previous section most 5' ends detected in our transcript end analysis are most likely generated post-transcriptionally. Since in plant mitochondria no 5' exonucleolytic activity has so far been detected these ends are most likely derived from endonucleolytic cleavage (9). To obtain experimental evidence for an endonucleolytic 5' processing reaction we used an experimental approach in which the mitochondrial 5S rRNA was used as an 'anchor' molecule. This rRNA was ligated to the 3' end of the potential 5' cleavage product (for details see Materials and Methods section and Figure 6A).

For the *cox1* 5' cleavage product a 195-bp cDNA fragment is expected in an amplification reaction with primers Atcox1-lm.H (-329 to -291) and At5S-mega.R (+107 to +82) on a cDNA template synthesized from

Table 2. 5' ends from mitochondrial mRNA from *A. thaliana*

Major ends			
Gene	Position relative to the start codon ^b (position in nc_001284)	Flanking sequences (5' to 3'), the most prominent ends are underlined with the most dominant end given in a large letter	Remarks
Transcript ends at conserved promoter motifs			
<i>atp8</i>	-157 (129 752)	AAACATCAATCTCATAAGA <u>G</u> AAGAAATCTCTATGCCCCCT ^d	CNM1, also detected in (12) ^c
<i>atp8</i>	-228/4 (129 681/5)	TTCCATACCATAACATAATAT <u>A</u> GAA <u>T</u> CGATTTCTTTCTGA ^d	CNM2, also detected in (12) ^c
<i>atp6-1</i>	-200 (111 550)	GCAGCCAATAATACGTATAT <u>A</u> GAAAGAGGACTGCTTACGGG ^d	CNM2, also detected in (12)
<i>ccmB (ccb2)</i>	-140 (30 323)	GTCTTTTAGTAAGCGTATAT <u>A</u> AGCAGCAGTTTAGTGTATAA ^d	CNM2, ^c
<i>cox2</i>	-205 (42 833)	GTTGGTTTCGTATATAAGAA <u>G</u> ACTGCTGCTTTTAGGAGAGT ^d	^a , CNM2, as -210 (+TAP ^c) and -205 (-TAP) detected in (12)
5' mRNA ends generated by RNase Z cleavage at the 3' ends of upstream located tRNAs			
<i>ccmF_C (ccb6c)</i>	-124 (53 735)	TTCAAGTCCCTCCTTCCGCT <u>C</u> TGGTGTTCGAACTAGTCAT	The 5' terminal cytidine overlaps with the 3' terminal nucleotide of tRNA ^{Gly}
<i>rps3-rpl16</i>	-163 (28 896)	TCAAGTCCCTGCTATACCCAA <u>A</u> CCTACCTTACCACACTATAT	The adenosine is one nucleotide downstream of the tRNA ^{Lys} 3' terminus. A processed end at-164 has been detected in (12)
5' mRNA ends most likely generated by RNase Z cleavage at the 3' ends of upstream located t-elements			
<i>rps4</i>	+2 (83 115)	TTTCGAGTCCCAGGGACGCA <u>A</u> TGTGGCTGCTTAAAAAAGTGA	t-element, RNase Z
<i>ccmF_{N1} (ccb6n1)</i>	-66 (233 108)	TTCGAATCCTGCCACCTTTC <u>T</u> TGTGGATCATCCTGTGGTTAC	t-element, RNase Z
<i>cox1</i>	-241 (351 654)	TTTCGAATCCAATGGTAGGCC <u>T</u> GGGCTCTCTACGCAGGAGA	t-element, RNase Z, also detected as processed end in (12)
5' ends directly adjacent to stem-loop structures			
<i>rpl5-cob</i>	-459 (57 315)	CTAACTAAATAGTTAGTTAA <u>G</u> AAGACGCGTGCTTTGTTGG	SL, RNase Z
<i>rpl5-cob</i>	-406 (57 368)	ATTTCACTATAAGATAAGAT <u>A</u> AGAGTGAAGTCTTTGTTTG	SL, RNase Z
<i>atp6-2</i>	-268 (298 137)	GTTTTGTTGGTCAACAAC <u>C</u> ACCACAACTTTATAGTTG	SL, RNase Z
<i>nad7</i>	-375 (131 696)	AAGCCCTCTCTTTATATAAA <u>G</u> AGGGCTTTCGCGCTTCGGAT	SL, RNase P
Miscellaneous transcript ends			
<i>atp9</i>	-84 (278 811)	CCGTGGGTATAAGATCGCAA <u>A</u> GAATGCATTCCAAGTGAGAT	Identical with the 5' terminus of the <i>nad6</i> mRNA, also detected as processed end at -83 by (12)
<i>nad6</i>	-179 (77 438)	CGTGGGTACAAGATCGAAAA <u>G</u> AATGCATTTAAATGGATGCAT	Identical with the 5' terminus of the <i>atp9</i> mRNA
<i>nad5</i>	-21 (143 019)	TGCTTTGATTTTAGGTTTTT <u>G</u> ATCACACTCGAAATTTAAAT	^a
<i>nad5</i>	-72 (143 070)	GTTCGCCGACGTCGAGAAA <u>T</u> CTATCACTCAATCACTGGCC	^a
<i>nad2</i>	-122 (81 419)	TTGAGCCGAATCACTATCAT <u>T</u> ATATTATATATTCATTTGTTG	
<i>ccmF_{N1} (ccb6n1)</i>	-149 (233 191)	AGCGCATCTGTTTCGGGT <u>A</u> CAGGGCAGCAGGGCGAAGCA	
<i>cox2</i>	-151 (42 779)	AACTCAAAATATTGTAAGAA <u>G</u> AAGAAGAATCTTACGCCCA	
<i>rpl2-mitB (orfX)</i>	-198 (154 546)	AAAATAATATAAATAATAAT <u>A</u> TATAGAAATATAGAAAAGAAAT	actually -194 (a four bp deletion in the 5' UTR of this gene was found in ecotype Col mtDNA)
<i>nad9</i>	-202 (24 437)	CCACGAATTTGATCATTCGC <u>T</u> TTCAATCAAATAGATTGTC	
<i>nad4</i>	-228 (161 465)	TCCGATAAAGAATGCAAAA <u>T</u> AGAAAGAATATAAATAAAGG	
<i>nad3-rps12</i>	-229 (261 235)	GAGGACCAC TAGTGGGGAAG <u>A</u> GTCCTTGGAGCGTCTTTATG	
<i>ccmB (ccb2)</i>	-347/6 (30 116/7)	GGCAAGGCATATAAGTAAGT <u>A</u> AAAGTTGGAAGCCTAAGCCAA	
<i>cox3</i>	-378 (217 902)	GAGTGGAGGGGAAGGTTTAA <u>T</u> ACCGAATTAGCAGGGATAGA	
<i>rps7</i>	-448 (314 179)	ACTCATATATATATATATAT <u>G</u> ATGACTTATATCTAAATGGA	
<i>ccmC (ccb3)</i>	-484/2 (241 242/0)	TATCGATCGGCGTAAGAAAC <u>G</u> ATCTACTATCTTTATTCCC	^a
<i>nad1</i>	-645 (319 035)	GAACCCCTGGCCATTAACCCG <u>A</u> AAATCTCGGGTTGGAGTCTTC	^a
<i>atp1</i>	-124 (303 813)	AAGTTAAGATAGATGTAACA <u>A</u> TTTGAACCAAAGGCCTTTT	
<i>atp1</i>	-267 (303 956)	GCTCAACTCGACATATGGAG <u>C</u> CAGGGGCTTGATCAGGAATT	
<i>atp1</i>	-361 (304 050)	CCGTGGGCGGGAAC TAGCCG <u>C</u> CCGCATCGATTGGGAGGTGC	

(continued)

Table 2. Continued

Gene	Position relative to the start codon ^b (position in nc_001284)	Flanking sequences (5' to 3'), the most prominent ends are underlined with the most dominant end given in a large letter	Remarks
Major ends			
<i>matR</i>		No clear 5' end detectable	
<i>ccm1F_{N2}</i> (<i>ccb6n2</i>)	-57 (257 533)/-35 (257 511)	<u>TTGGCGAAATCATGTGGGGC</u> <u>TT</u> CAGAAGAGTGTCCCTACGAC	
<i>nad4L-atp4</i> (<i>orf25</i>)	-318 (189 549)/-294 (189 525)	<u>CGTAAATCGGGCAAAAGA</u> GGGGAAACCGGGGAAATCA <u>C</u> TAGATCCCTCATGGGATCAAA AA <u>C</u> AACCTATACCCGAAGTCCCA	
Minor ends			
<i>atp1</i>	-1898 (305 587)	AAACTATCAATTTCAATAAGA <u>G</u> AAGAAAGAAAGTAGCCTAGT ^d	CNMI, also detected in (12)
<i>atp6-2</i>	-63 (297 332)/-44 (297 913)	TATAAGTAGACACCAATAAT <u>T</u> ACCACCCCTTCTAGTGT <u>C</u> GGGT ACAGTAGCTCTCGCAG	
<i>rpl2-mitB</i> (<i>orfX</i>)	-119 (154 625)	CTACTAAAGTAAAGTTGAC <u>C</u> ACCCGGACCAGACAGAGGTGGCCG	
<i>nad7</i>	-68 (132 003)	GTCTCGAAGGAGATGCAT <u>T</u> CTGGTACAAGTGGTATTGGA	^a
<i>nad1</i>	-109 (131 962)	CAAGCAAGCGTAAATAAGGG <u>A</u> GGGGACTAGGGGTGGAAGG	^a
<i>nad1</i>	-149 (318 539)	CTTTATTCGGGAGTAAAGT <u>G</u> ATAGGGGTTCCGTCATTTG	^a
<i>nad1</i>	-355 (318 745)	TCAAGAGGCTTTTTCGACT <u>T</u> CTTGAACGAGAGGGCGGACG	^a

^aExact location possibly imprecise due to superposed signals in the sequencing chromatogram.

^bThe first nucleotide upstream of the start codon corresponds to position -1.

^cExact location of the 5' end determined by analysis of primer extension products separated on a high-resolution polyacrylamide gel beside a sequencing reaction with the same primer.

^dConserved promoter motifs are given in italics. ^eTAP: tobacco acid pyrophosphatase.

26S rRNA: GTACAAGATCGAAAAGAATGCATT
nad6: GTACAAGATCGAAAAGAATGCATT
atp9: GTA*t*AAGATCGcAAAGAATGCATT

Figure 5. Sequences found conserved at the 5' termini of the 26S rRNA as well as *atp9* and *nad6* mRNAs. The 5' terminal nucleotides are underlined. Differing nucleotide identities are indicated with lower case letters in italics.

oligonucleotide At5S-5 (+118 to +100). When products of this PCR were separated on an agarose gel, the predominant cDNA fragment indeed showed the expected size (Figure 6B). This product was thus cloned and sequenced. In a total number of 20 clones, 19 cDNA inserts contained 3' ends corresponding to nucleotide position -242. This is the position exactly upstream of the major mature 5' end of the *cox1* mRNA at -241 (Figures 2D and 3, Tables 2 and 3). An additional end was found at -239 for which the corresponding 5' end of the *cox1* mRNA at -238 has not been detected, indicating that the ends at these positions might vary slightly more than seen in the above described analysis of the *cox1* 5' mRNA extremities (Figures 2 and 3). The results from this approach strongly suggest that the major 5' end of the *cox1* transcript around position -241 is generated by endonucleolytic cleavage. Interestingly 13 cDNA clones of the *cox1* 5' leader RNAs contained non-encoded cytidines and adenosines at the 3' termini. These nucleotides might be added by a terminal tRNA nucleotidyltransferase, which adds the 5' -CCA-3' triplet to tRNAs and indeed a perfectly conserved pseudo-uridine arm is found upstream of this 3' end. In fact, 74 nucleotides of the *cox1* 5' leader molecule can be folded into a tRNA-like structure (Figure 6C), a so-called t-element (29). This clearly indicates that the 5' end of the *cox1* mRNA at position -241 is generated by a cleavage reaction most likely catalyzed by RNase Z. This result raises the question whether also other termini of mitochondrial mRNAs might be generated by RNase Z or by RNase P.

But also mRNA ends not associated with a t-element are generated by an endonucleolytic cleavage. Analogous experiments performed to detect potential processed *atp9* 5' leader molecules revealed that also the major 5' end at position -83 is derived from an endonucleolytic cut (Supplementary Figure 28).

t-elements contribute to the formation of 5' and 3' ends

To check whether other mitochondrial 5' and also 3' mRNA ends could be derived from cleavage by RNase Z or RNase P we screened the complete sequence of the *A. thaliana* mitochondrial DNA for the presence of the highly conserved 5'-GGTTCRANYCC-3' motif of the pseudo-uridine arm of tRNAs. Apart from the canonical tRNAs this motif is present upstream of the 5' ends of *rps4* and *ccmF_{N1}* as well as downstream of the gene-internal 3' ends of *ccmC* and *nad6*, indicating that these conserved sequences could be involved in the generation of 5' ends as well as 3' termini. Folding of these sequences revealed

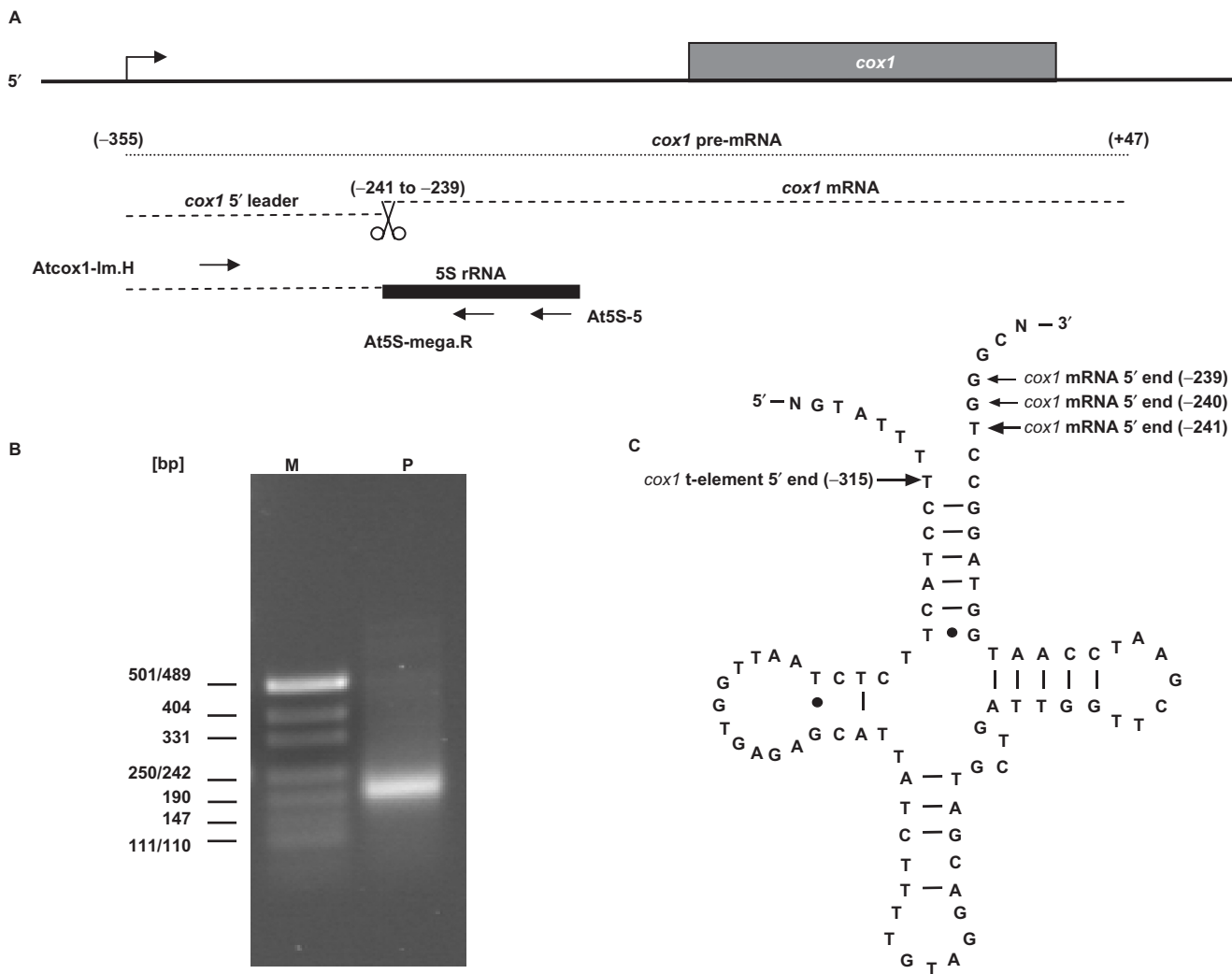


Figure 6. The 5' terminus of the *cox1* steady state mRNA is generated by an endonucleolytic cleavage. (A) The *cox1* gene (gray box) is transcribed from a promoter (bent arrow) located 355 nucleotides upstream of the ATG. The precursor RNA (dotted line) is processed by endonucleolytic cleavage (scissor) into a mature *cox1* mRNA and a small 5' leader RNA. In an RNA ligation reaction (bottom line) endogenous mitochondrial 5S rRNA (black bar) is ligated to the *cox1* 5' leader RNA and then used as anchor in the cDNA synthesis initiated at primer At5S-5 and a subsequent PCR with primer At5S-mega.R and Atcox1-lm.H. (B) PCR products (lane P) obtained in the reaction described earlier are separated on an agarose gel alongside DNA marker fragments (lane M). A single 200-bp product is obtained which contains the 5S rRNA ligated to the 3' end of the potential *cox1* 5' leader fragment. (C) The leader sequence upstream of the mature *cox1* 5' end can be folded into a tRNA-like structure forming a perfectly conserved ψ -uridine arm. The 5' end of this t-element was also mapped in a CR-RT-PCR analysis using 5S rRNA as anchor (Supplementary Figure 31).

that the conserved motifs are parts of t-elements with ends coinciding with the experimentally mapped transcript termini (Figure 7).

To gain more experimental evidence for the importance of these t-elements in RNA processing, we searched for 3' ends mapping downstream of the reading frame-internal major *nad6* and *ccmC* 3' termini. Such ends would be expected if the t-elements are cleaved off from respective precursor RNAs as postulated. For the identification of the 5' and 3' ends of the *nad6* t-element we used the same approach as described above for the detection of the 3' end of the *cox1* 5' leader. To identify the extremities of the *ccmC* t-element, 18S rRNA was used as anchor molecule instead of 5S rRNA in an otherwise identical approach (Supplementary Figures 29 and 30).

Direct sequencing of the PCR products revealed 5' and 3' ends of the *ccmC* t-element at position -36 (240021) and +37 (239951), consistent with the t-element model presented in Figure 7. The t-elements contain single non-encoded cytidines at their 3' termini. An experimental approach to detect the +37 end at the *ccmC* mRNA failed. Obviously t-element processing is very efficient, so that this 3' end cannot be considered as a genuine *ccmC* mRNA terminus.

In the case of the *nad6* t-element, a 5' terminus is found two nucleotides downstream (-15) of the reading frame-internal mRNA 3' end (-17) strongly suggesting a cleavage by RNase P at the site predicted by the model (Figure 7). Likewise, the 3' terminus of this t-element was mapped at position +40 also consistent with the predicted t-element secondary structure and a cleavage by RNase Z

Table 3. 3' ends of *cox1* 5' leader RNA molecules

Clone	3' end position in the mitochondrial genome	3' end position in respect to the <u>ATG</u> (+1)	Non-encoded nucleotides
#27	351 656/351 655 ^a	-243/-242 ^a	
#6	351 655	-242	
#8	351 655	-242	
#11	351 655	-242	
#28	351 655	-242	
#29	351 655	-242	
#9	351 655	-242	A
#18	351 655	-242	A
#5	351 655	-242	C
#7	351 655	-242	C
#13	351 655	-242	C
#20	351 655	-242	C
#1	351 655	-242	CC
#2	351 655	-242	CC
#14	351 655	-242	CC
#16	351 655	-242	CA
#17	351 655	-242	CAAAA
#22	351 655	-242	CAAAA
#19	351 655	-242	CCAAA
#10	351 652	-239	

^aA single nucleotide at the ligation site could be either assigned to *cox1* or 5S rRNA. Thus, the *cox1* t-element ends at position -243 or -242.

(Figure 7). Like the *cox1* t-element derived cDNAs, also cDNA clones originating from the *nad6* t-element contained non-encoded cytidines and adenosines (Supplementary Figure 29K).

In summary, these results suggest that both tRNA processing enzymes generate 5' and 3' ends of plant mitochondrial mRNAs. This does not only occur at canonical tRNAs such as tRNA^{Gly} and tRNA^{Lys} located upstream of *ccmF_C* and *rps3* and tRNA^{Ser} encoded downstream of the *atp6* genes, but also at t-elements. RNase Z and RNase P are thus responsible for endonucleolytic processing of at least a subset of plant mitochondrial mRNAs.

DISCUSSION

General features of plant mitochondrial mRNAs

In our study, we have analyzed the 5' and 3' mRNA ends of all annotated mitochondrial protein-coding genes in *A. thaliana*. This comprehensive study reveals several general features of these organellar transcripts in plants. First, almost all transcripts have a single 3' end that in the vast majority of mRNAs scatters over only few nucleotides. The only exception is the *ccmF_C* mRNA, which has two 3' ends 105 nucleotides apart. Second, for many genes we find mRNAs with different 5' ends. Third, with the exception of the primary 5' ends within promoter sequences, there is no conserved sequence motif evident at the 5' termini or at the 3' ends. Thus other determinants govern the post-transcriptional generation of these ends. Fourth, the majority of the 5' ends are most likely derived from processing. There are several arguments that support this conclusion. Most of the ends identified do not terminate in any known promoter motif (12). In addition,

upstream of many of the 5' ends detected in our study promoter sequences have been found for instance in the case of *atp9*, *atp1*, *cox1*, *cox2*, *rps3* (12) and potential conserved promoters can be predicted upstream of *cox3* (-640 in ecotype C24, here 5' ends have been found at this position), *nad3* (-1320 and 1230), *nad1* (-1400) and *ccmF_{N1}* (-890). Moreover, our analysis shows that the 5' ends of the *cox1* and the *atp9* mRNAs are generated by endonucleolytic cleavage. Analogous cleavages may generate other secondary 5' and 3' ends (for a detailed discussion see subsequently).

Our analysis also revealed non-encoded nucleotides most likely attached to the 3' ends. Mostly only single adenosines are found and less frequently cytidines, only in rare instances thymidines (corresponding to uridines in the RNA) or guanosines. The extensions can be grouped into two categories as observed previously (18,22,23,46). They are either oligohomopolymeric adenosine stretches (up to 24 adenosines) or short extensions of adenosines and cytosines. However, the vast majority of the mRNAs do not have non-encoded nucleotides consistent with previous assumptions (9).

Our analyses also confirm that RNA editing in the 5' and 3' UTRs is a rare event. Apart from the editing sites found previously (47), we detected new editing sites only in the 5' UTR of *ccmC*. However, a blast search revealed that these editing sites are located within a 186-bp duplication of a part of the *nad6* reading frame, where exactly the same cytidines are edited. The three C to U conversions are found at positions -286 (position 241 044 = 77069 in *nad6*), -308 (241 066 = 77091) and -316 (241 074 = 77099). We also identified a new editing site in the *atp1* reading frame at position 302 265, which is only edited at a rate of about 35%. In contrast, some postulated editing sites were never seen edited in our analysis. This includes cytidines at position 219 244 in the pseudo *sdh4* gene and at position 260 938 in *nad3*.

Taken together, this comprehensive investigation of mitochondrial mRNA ends in *A. thaliana* provides new general insights into the nature and the generation of these ends and thus represents a broad basis for further detailed studies of individual mRNA ends.

Endonucleolytic generation of both 5' and 3' ends

As mentioned earlier most or at least a large portion of the 5' termini could be derived from post-transcriptional processing. We provide clear experimental evidence for an endonucleolytic generation of four 5' mRNA ends (Table 4). While this kind of processing was expected for *ccmF_C* and *rps3*, where tRNAs are located upstream of the genes, we now find that also a t-element upstream of *cox1* acts as *cis*-element most likely directing cleavage by RNase Z. A similar scenario can be expected for the 5' ends of *ccmF_{N1}* and *rps4*, where also t-elements can be found 5' of the mature mRNA ends.

In line with the role of such elements in 5' end formation, it can be assumed that even minimal secondary structures such as simple stem-loops can be recognized as processing signals by tRNA processing enzymes. Three such structures might direct RNase Z to generate

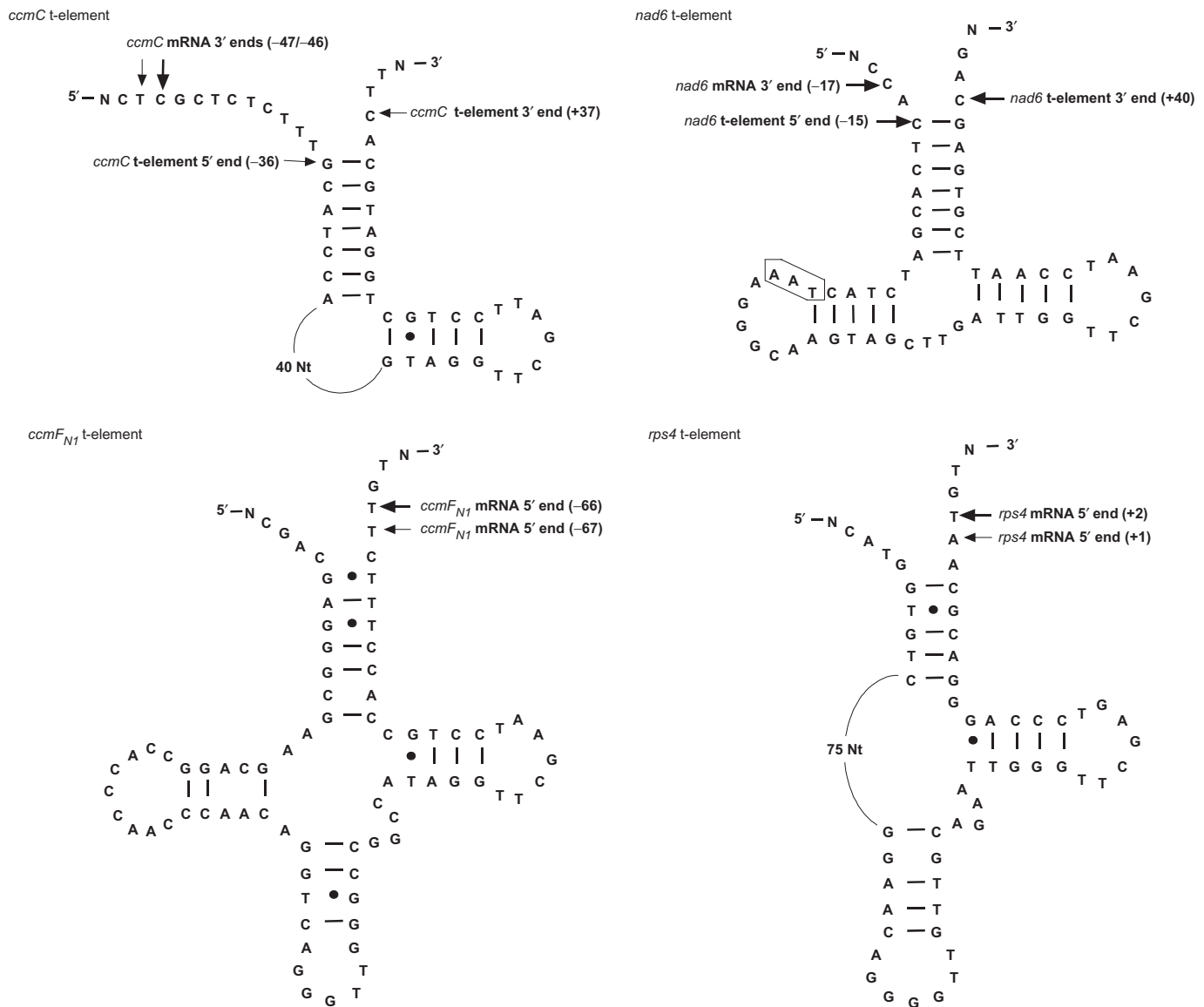


Figure 7. In the *A. thaliana* mitochondrial genome several t-elements can be identified. Either the 5' or 3' ends of these t-elements coincide with 5' and/or 3' termini of mitochondrial mRNAs. This suggests that these t-elements direct cleavage of the respective precursor RNAs by RNase Z and RNase P *in vivo*.

Table 4. Transcript ends associated with tRNAs, t-elements or stem-loop structures

Gene	Nature of mRNA end	Position of mRNA end	Kind of secondary structure	Endoribonuclease involved	Evidence for endonucleolytic generation
<i>ccmF_C</i> (<i>ccb6c</i>)	5'	-124	tRNA ^{Gly}	RNase Z	hypothetical
<i>rps3</i>	5'	-163	tRNA ^{Lys}	RNase Z	hypothetical
<i>rps4</i>	5'	+2	t-element	RNase Z	hypothetical
<i>ccmF_{N1}</i> (<i>ccb6n1</i>)	5'	-66	t-element	RNase Z	hypothetical
<i>cox1</i>	5'	-241	t-element	RNase Z	experimental
<i>rpl5</i>	5'	-459	acceptor-stem-like stem-loop	RNase Z	hypothetical
<i>rpl5</i>	5'	-406	acceptor-stem-like stem-loop	RNase Z	hypothetical
<i>atp6-2</i>	5'	-268	acceptor-stem-like stem-loop	RNase Z	hypothetical
<i>nad7</i>	5'	-375	acceptor-stem-like stem-loop	RNase P	hypothetical
<i>atp6-1</i>	3'	+45	tRNA ^{Ser}	RNase P	hypothetical
<i>atp6-2</i>	3'	+45	tRNA ^{Ser}	RNase P	hypothetical
<i>ccmC</i> (<i>ccb3</i>)	3'	-46	t-element	RNase P	experimental ^a
<i>nad6</i>	3'	-17	t-element	RNase P	experimental ^a

^amRNA 3' end generated most likely by concerted action of endo- and exonucleolytic activities.

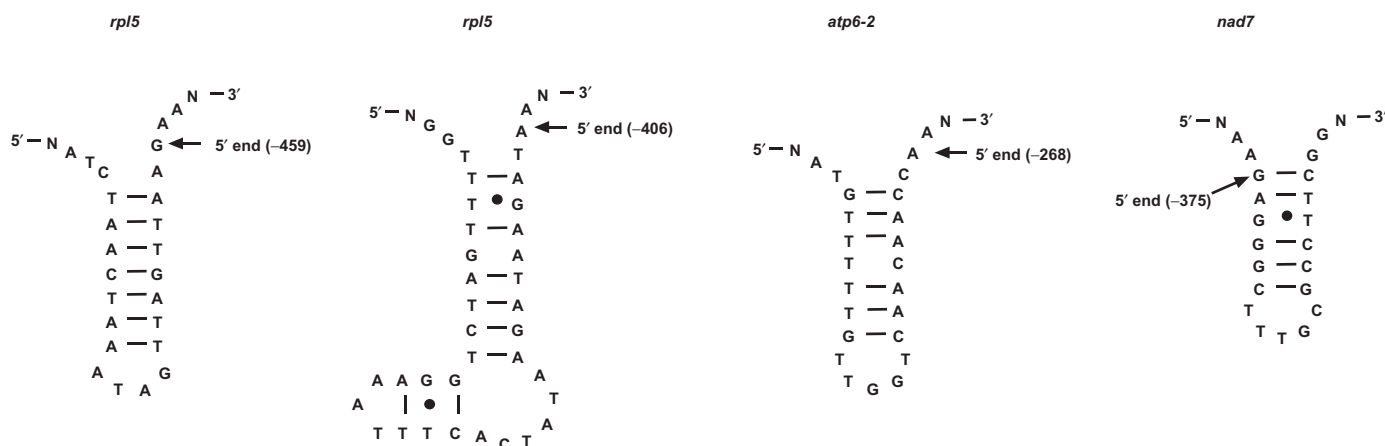


Figure 8. Stem-loop structures found at 5' mRNA ends. Like the t-elements the ends of the stem-loop structures coincide with mRNA termini. These stem-loop structures might thus mimic tRNA acceptor stems and direct cleavage by RNase Z or RNase P.

two *rpl5* 5' termini (-459 and -406) and an *atp6-2* 5' end (-268), while a simple SL downstream of the *nad7* (-375) 5' terminus suggests RNase P to generate this end (Figure 8). But there must be an additional mode of processing since no t-element or any other obvious secondary element is found at the 5' cleavage site of *atp9* mRNA and many other potential cleavage sites. Maybe a totally different set of proteins including a so far unknown endonuclease generates these ends. But it might be also possible that specific binding proteins allow an action of the tRNA processing enzymes on substrates that do not resemble a tRNA or parts of the tRNA secondary structure.

While exonucleolytic processes had been suggested for formation of at least some 3' ends, we now found evidence for direct endonucleolytic generation of 3' ends or at least for the involvement of endonucleolytic cleavage in 3' end formation. This has been similarly found during the investigation of Ogura cytoplasmic male sterility (CMS) in Brassica hybrids. Here a t-element was identified downstream of the *orf138*, which is responsible for the CMS phenotype. Most likely cleavage by RNase P generates the 3' end of *orf138* mRNAs. In addition also 3' cleavage of the t-element is observed giving rise to stable t-element RNA (30).

Analogous to 5' end generation, direct endonucleolytic 3' end generation is expected when a canonical tRNA is located 3' to an mRNA, which results in a cleavage by RNase P as for instance for both *atp6* mRNAs (Table 4). At least an essential role of endonucleolytic cleavage in 3' end formation can be assumed for *nad6* and *ccmC* mRNAs. For the former we found the 3' end of the mRNA and the 5' end of the respective t-element to be separated by a single nucleotide, indicating that this adenosine might be removed in a trimming process involving RNR1. Likewise, we identified the 5' terminus of the *ccmC* t-element 10 nucleotides downstream the respective 3' mRNA ends suggesting these nucleotides also to be removed by RNR1. It would thus be interesting to see whether such a concerted action of endonucleolytic cleavage and exonucleolytic trimming is frequently observed in the formation of mitochondrial

mRNA ends. Interestingly, both ends derived from this kind of processing are located within the reading frame. Also in analogy with 5' end processing, it is also possible that stem-loops might be *cis*-elements recognized by tRNA processing enzymes resulting in an endonucleolytic generation of a 3' end. Such a scenario might be expected for the 3' end formation of the *cox2* mRNA directed by a SL (Figure 4), but *in vitro* studies are required to discriminate between endo- and/or exonucleolytic generation of this end.

The involvement of tRNA maturation in dismantling large precursor RNAs occurs in mitochondria of animals (48). Here protein-coding genes are usually interspersed by at least a single tRNA. After transcription of a complete strand of the mitochondrial genome endonucleolytic removal of the tRNAs results in the generation of individual mRNAs. This schema of processing seems at least partially conserved in plant mitochondria although the protein coding genes are often separated by large spacer sequences.

Function of secondary 5' ends

The detection of many secondary 5' ends raises the question why some mRNAs retain their original primary 5' ends while most others are processed. A size reduction of the 5' UTRs would be a simple explanation, but even mRNAs with short original 5' UTR (for instance *atp9*) are processed, while some of the processed ends (*cox3*, *rpl5*) define larger 5' UTRs.

Of course the nature of the 5' UTRs could influence translation and it is possible that some mRNAs are only accessible for translation after 5' processing. Here two scenarios are possible. Either a potential ribosome entry sequence is blocked in the non-processed RNA or binding of specific translation factors is not possible prior to processing. The first scenario seems to be rather unlikely since no Shine-Dalgarno-like sequence has been identified in plant mitochondria. Although conserved sequence elements have been found in some mRNAs (49) the restriction of this motif to certain mRNAs excludes a general function. Moreover, there is

so far no experimental evidence for the importance of these conserved sequences. More likely seems the second scenario. mRNA-specific translation factors have indeed been found in mitochondria from *Saccharomyces cerevisiae* (50) and in chloroplasts from *Chlamydomonas reinhardtii* (51) and higher plants (52). In the latter case, the CRP1 protein was found to specifically interact with 5' UTRs of *petA* and *psaC* mRNAs and this protein activates the translation of these RNAs. CRP1 belongs to the PPR protein family of which many members are transported to mitochondria, where they could exhibit similar functions.

Is there a general structure for RNA-stabilizing elements in plant mitochondria?

In *A. thaliana* there seems to be only little variability at the 3' ends of mitochondrial mRNAs and apart from a slight scattering over a few nucleotides the termini are determined very clearly. This provokes two questions: how are 3' ends generated and defined and what mechanisms exist that stabilize mRNAs and prevent them from degradation? Both questions might be closely related since exonucleolytic 3' end maturation can be considered as a process leading from a precursor RNA towards as stable mature mRNA.

Recently, two 3' exoribonucleases have been identified as important *trans*-factors necessary for the generation of both *atp9* and *atp8* mRNAs (21) with the PNPase also being responsible for the degradation of dispensable RNAs (53). An important role of these exoribonucleases in 3' end processing is highly likely and protection of an RNA from exonucleolytic degradation is probably crucial for stability of a plant mitochondrial mRNA.

In chloroplasts, SLs located in the 3' UTRs, are important for both 3' end trimming and stability of the mRNA (54). Likewise, such structures have been suggested to protect plant mitochondrial mRNAs in some instances (17,55) but the comprehensive survey now clearly reveals that the presence of such structures at the 3' end of mitochondrial mRNA is an exception rather than the rule. The vast majority of the mRNA 3' UTRs does not contain any obvious single or double SLs suggesting that other *cis*-elements determine the stability of an mRNA. Thus, in plant mitochondria there seems to be an mRNA stabilizing and also 3' processing system that is different from the one that has been suggested for chloroplasts. This discrepancy, obvious at least in terms of *cis*-elements, is somehow surprising since both organelles contain very similar and identical, respectively, 3' exoribonucleases. Thus one might speculate that different auxiliary factors might communicate between the different *cis*-elements and the exoribonucleases.

But what determines 3' end formation and mRNA stability in plant mitochondria? Our data clearly show that there is neither a conserved primary structure nor a conserved secondary structure element. Thus, individual mRNA-specific determinants might exist, either at the primary sequence or secondary structure level and it might well be that also individual *trans*-factors might be involved.

This might be also reflected by the ribonucleases involved in 3' end formation. While some ends might be generated exclusively by exoribonucleases, we now found clear evidence for the participation of endonucleolytic activities in this process. This is apparent for the two *atp6* mRNAs, whose 3' ends are directly created by the cleavage of RNase P at the 5' end of tRNA^{Ser} (UGA). But also the formation of the 3' ends of the truncated *nad6* and *cemC* mRNAs requires endonucleolytic cuts. However, these 3' termini are formed by a concerted action of both endo- and exoribonucleases.

To elucidate the mechanism(s) of mRNA 3' end formation and mRNA stability in plant mitochondria more studies are required and the data presented here provide a fundamental basis for future functional analyses.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

This work was supported by the DFG (Bi 590/6-1 and 6-2) and by a fellowship of the Studienstiftung des deutschen Volkes to J.F. The authors are very grateful to C. Guha for technical assistance and thank all co-workers and students who directly or indirectly contributed to this project. Funding to pay the Open Access publication charges for this article was provided by the DFG.

Conflict of interest statement. None declared.

REFERENCES

- Clifton, S.W., Minx, P., Fauron, C.M., Gibson, M., Allen, J.O., Sun, H., Thompson, M., Barbazuk, W.B., Kanuganti, S. *et al.* (2004) Sequence and comparative analysis of the maize NB mitochondrial genome. *Plant Physiol.*, **136**, 3486–3503.
- Handa, H. (2003) The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic Acids Res.*, **31**, 5907–5916.
- Kubo, T., Nishizawa, S., Sugawara, A., Itchoda, N., Estiati, A. and Mikami, T. (2000) The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNA(Cys)(GCA). *Nucleic Acids Res.*, **28**, 2571–2576.
- Notsu, Y., Masood, S., Nishikawa, T., Kubo, N., Akiduki, G., Nakazono, M., Hirai, A. and Kadowaki, K. (2002) The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. *Mol. Genet. Genomics*, **268**, 434–445.
- Ogihara, Y., Yamazaki, Y., Murai, K., Kanno, A., Terachi, T., Shiina, T., Miyashita, N., Nasuda, S., Nakamura, C. *et al.* (2005) Structural dynamics of cereal mitochondrial genomes as revealed by complete nucleotide sequencing of the wheat mitochondrial genome. *Nucleic Acids Res.*, **33**, 6235–6250.
- Satoh, M., Kubo, T., Nishizawa, S., Estiati, A., Itchoda, N. and Mikami, T. (2004) The cytoplasmic male-sterile type and normal type mitochondrial genomes of sugar beet share the same complement of genes of known function but differ in the content of expressed ORFs. *Mol. Genet. Genomics*, **272**, 247–256.
- Sugiyama, Y., Watase, Y., Nagase, M., Makita, N., Yagura, S., Hirai, A. and Sugiura, M. (2005) The complete nucleotide sequence

- and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. *Mol. Genet. Genomics*, **272**, 603–615.
8. Unseld, M., Marienfeld, J.R., Brandt, P. and Brennicke, A. (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nat. Genet.*, **15**, 57–61.
 9. Gagliardi, D. and Binder, S. (2007) Expression of the plant mitochondrial genome. In Logan, D. (ed.), *Plant mitochondria*. Blackwell Publishing, Ames, IA, USA, pp. 50–96.
 10. Dombrowski, S., Hoffmann, M., Kuhn, J., Brennicke, A. and Binder, S. (1998) On mitochondrial promoters in *Arabidopsis thaliana* and other flowering plants. In Möller, I.M., Gardeström, P., Glimelius, K. and Glaser, E. (eds), *Plant Mitochondria: From Gene to Function*. Backhuys Publications, Leiden, Netherlands, pp. 165–170.
 11. Giese, A., Thalheim, C., Brennicke, A. and Binder, S. (1996) Correlation of nonanucleotide motifs with transcript initiation of 18S rRNA genes in mitochondria of pea, potato and *Arabidopsis*. *Mol. Gen. Genet.*, **252**, 429–436.
 12. Kühn, K., Weihe, A. and Börner, T. (2005) Multiple promoters are a common feature of mitochondrial genes in *Arabidopsis*. *Nucleic Acids Res.*, **33**, 337–346.
 13. Mulligan, R.M., Leon, P. and Walbot, V. (1991) Transcriptional and posttranscriptional regulation of maize mitochondrial gene expression. *Mol. Cell. Biol.*, **11**, 533–543.
 14. Finnegan, P.M. and Brown, G.G. (1990) Transcriptional and Post-Transcriptional Regulation of RNA Levels in Maize Mitochondria. *Plant Cell*, **2**, 71–83.
 15. Leino, M., Landgren, M. and Glimelius, K. (2005) Alloplasmic effects on mitochondrial transcriptional activity and RNA turnover result in accumulated transcripts of *Arabidopsis* orfs in cytoplasmic male-sterile *Brassica napus*. *Plant J.*, **42**, 469–480.
 16. Giegé, P., Hoffmann, M., Binder, S. and Brennicke, A. (2000) RNA degradation buffers asymmetries of transcription in *Arabidopsis* mitochondria. *EMBO Rep.*, **1**, 164–170.
 17. Dombrowski, S., Brennicke, A. and Binder, S. (1997) 3'-Inverted repeats in plant mitochondrial mRNAs are processing signals rather than transcription terminators. *EMBO J.*, **16**, 5069–5076.
 18. Kuhn, J., Tengler, U. and Binder, S. (2001) Transcript lifetime is balanced between stabilizing stem-loop structures and degradation-promoting polyadenylation in plant mitochondria. *Mol. Cell. Biol.*, **21**, 731–742.
 19. Kaleikau, E.K., Andre, C.P. and Walbot, V. (1992) Structure and expression of the rice mitochondrial apocytochrome b gene (*cob-1*) and pseudogene (*cob-2*). *Curr. Genet.*, **22**, 463–470.
 20. Saalau, E., Litvak, S. and Araya, A. (1990) The apocytochrome from alloplasmic line wheat (*T. aestivum*, cytoplasm-T. timopheevi) exists in two differently expressed forms. *Plant Sci.*, **66**, 237–246.
 21. Perrin, R., Meyer, E.H., Zaepfel, M., Kim, Y.J., Mache, R., Grienberger, J.M., Gualberto, J.M. and Gagliardi, D. (2004) Two exoribonucleases act sequentially to process mature 3'-ends of atp9 mRNAs in *Arabidopsis* mitochondria. *J. Biol. Chem.*, **279**, 25440–25446.
 22. Gagliardi, D. and Leaver, C.J. (1999) Polyadenylation accelerates the degradation of the mitochondrial mRNA associated with cytoplasmic male sterility in sunflower. *EMBO J.*, **18**, 3757–3766.
 23. Lupold, D.S., Caoile, A.G. and Stern, D.B. (1999) Polyadenylation occurs at multiple sites in maize mitochondrial *cox2* mRNA and is independent of editing status. *Plant Cell*, **11**, 1565–1578.
 24. Kunzmann, A., Brennicke, A. and Marchfelder, A. (1998) 5' end maturation and RNA editing have to precede tRNA 3' processing in plant mitochondria. *Proc. Natl Acad. Sci. USA*, **95**, 108–113.
 25. Hanic-Joyce, P.J. and Gray, M.W. (1990) Processing of transfer RNA precursors in a wheat mitochondrial extract. *J. Biol. Chem.*, **265**, 13782–13791.
 26. Marchfelder, A. (1995) Plant mitochondrial RNase P. *Mol. Biol. Rep.*, **22**, 151–156.
 27. Vogel, A., Schilling, O., Spath, B. and Marchfelder, A. (2005) The tRNase Z family of proteins: physiological functions, substrate specificity and structural properties. *Biol. Chem.*, **386**, 1253–1264.
 28. Frank, D.N. and Pace, N.R. (1998) Ribonuclease P: unity and diversity in a tRNA processing ribozyme. *Annu. Rev. Biochem.*, **67**, 153–180.
 29. Hanic-Joyce, P.J., Spencer, D.F. and Gray, M.W. (1990) *In vitro* processing of transcripts containing novel tRNA-like sequences ('t-elements') encoded by wheat mitochondrial DNA. *Plant Mol. Biol.*, **15**, 551–559.
 30. Bellaoui, M., Pelletier, G. and Budar, F. (1997) The steady-state level of mRNA from the *Ogura* cytoplasmic male sterility locus in *Brassica* cybrids is determined post-transcriptionally by its 3' region. *EMBO J.*, **16**, 5057–5068.
 31. Forner, J., Weber, B., Wietholter, C., Meyer, R.C. and Binder, S. (2005) Distant sequences determine 5' end formation of *cox3* transcripts in *Arabidopsis thaliana* ecotype C24. *Nucleic Acids Res.*, **33**, 4673–4682.
 32. Klein, M., Binder, S. and Brennicke, A. (1998) Purification of mitochondria from *Arabidopsis*. *Methods Mol. Biol.*, **82**, 49–53.
 33. Binder, S. (1995) Mitochondrial nucleic acid purification and analysis. *Methods Mol. Biol.*, **49**, 383–389.
 34. Kuhn, J. and Binder, S. (2002) RT-PCR analysis of 5' to 3'-end-ligated mRNAs identifies the extremities of *cox2* transcripts in pea mitochondria. *Nucleic Acids Res.*, **30**, 439–446.
 35. Sambrook, J. and Russel, D.W. (eds). (2001) *Molecular Cloning: A Laboratory Manual*, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
 36. McGinnis, S. and Madden, T.L. (2004) BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res.*, **32**, W20–25.
 37. Raczynska, K.D., Le Ret, M., Rurek, M., Bonnard, G., Augustyniak, H. and Gualberto, J.M. (2006) Plant mitochondrial genes can be expressed from mRNAs lacking stop codons. *FEBS Lett.*, **580**, 5641–5646.
 38. Giege, P., Knoop, V. and Brennicke, A. (1998) Complex II subunit 4 (*sdh4*) homologous sequences in plant mitochondrial genomes. *Curr. Genet.*, **34**, 313–317.
 39. Marker, C., Zemmann, A., Terhorst, T., Kiefmann, M., Kastenmayer, J.P., Green, P., Bachellerie, J.P., Brosius, J. and Huttenhofer, A. (2002) Experimental RNomics: identification of 140 candidates for small non-messenger RNAs in the plant *Arabidopsis thaliana*. *Curr. Biol.*, **12**, 2002–2013.
 40. England, T.E., Gumpert, R.I. and Uhlenbeck, O.C. (1977) Dinucleoside pyrophosphate are substrates for T4-induced RNA ligase. *Proc. Natl Acad. Sci. USA*, **74**, 4839–4842.
 41. Sugino, A., Snoper, T.J. and Cozzarelli, N.R. (1977) Bacteriophage T4 RNA ligase. Reaction intermediates and interaction of substrates. *J. Biol. Chem.*, **252**, 1732–1738.
 42. Binder, S., Hatzack, F. and Brennicke, A. (1995) A novel pea mitochondrial *in vitro* transcription system recognizes homologous and heterologous mRNA and tRNA promoters. *J. Biol. Chem.*, **270**, 22182–22189.
 43. Binder, S. and Brennicke, A. (1993) Transcription initiation sites in mitochondria of *Oenothera berteriana*. *J. Biol. Chem.*, **268**, 7849–7855.
 44. Hoffmann, M. and Binder, S. (2002) Functional importance of nucleotide identities within the pea *atp9* mitochondrial promoter sequence. *J. Mol. Biol.*, **320**, 943–950.
 45. Brown, G.G., Auchincloss, A.H., Covello, P.S., Gray, M.W., Menassa, R. and Singh, M. (1991) Characterization of transcription initiation sites on the soybean mitochondrial genome allows identification of a transcription-associated sequence motif. *Mol. Gen. Genet.*, **228**, 345–355.
 46. Williams, M.A., Johzuka, Y. and Mulligan, R.M. (2000) Addition of non-genomically encoded nucleotides to the 3'-terminus of maize mitochondrial mRNAs: truncated *rps12* mRNAs frequently terminate with CCA. *Nucleic Acids Res.*, **28**, 4444–4451.
 47. Giege, P. and Brennicke, A. (1999) RNA editing in *Arabidopsis* mitochondria effects 441 C to U changes in ORFs. *Proc. Natl Acad. Sci. USA*, **96**, 15324–15329.
 48. Ojala, D., Montoya, J. and Attardi, G. (1981) tRNA punctuation model of RNA processing in human mitochondria. *Nature*, **290**, 470–474.
 49. Pring, D.R., Mullen, J.A. and Kempken, F. (1992) Conserved sequence blocks 5' to start codons of plant mitochondrial genes. *Plant Mol. Biol.*, **19**, 313–317.
 50. Naithani, S., Saracco, S.A., Butler, C.A. and Fox, T.D. (2003) Interactions among COX1, COX2, and COX3 mRNA-specific translational activator proteins on the inner surface of the

- mitochondrial inner membrane of *Saccharomyces cerevisiae*. *Mol. Biol. Cell*, **14**, 324–333.
51. Auchincloss, A.H., Zerges, W., Perron, K., Girard-Bascou, J. and Rochaix, J.D. (2002) Characterization of Tbc2, a nucleus-encoded factor specifically required for translation of the chloroplast psbC mRNA in *Chlamydomonas reinhardtii*. *J. Cell Biol.*, **157**, 953–962.
52. Schmitz-Linneweber, C., Williams-Carrier, R. and Barkan, A. (2005) RNA immunoprecipitation and microarray analysis show a chloroplast Pentatricopeptide repeat protein to be associated with the 5' region of mRNAs whose translation it activates. *Plant Cell*, **17**, 2791–2804.
53. Holec, S., Lange, H., Kuhn, K., Alioua, M., Borner, T. and Gagliardi, D. (2006) Relaxed transcription in *Arabidopsis* mitochondria is counterbalanced by RNA stability control mediated by polyadenylation and polynucleotide phosphorylase. *Mol. Cell. Biol.*, **26**, 2869–2876.
54. Hayes, R., Kudla, J. and Grussem, W. (1999) Degrading chloroplast mRNA: the role of polyadenylation. *Trends Biochem. Sci.*, **24**, 199–202.
55. Morikami, A. and Nakamura, K. (1993) Transcript map of oppositely oriented pea mitochondrial genes encoding the alpha-subunit and the subunit 9 of F₁F₀-ATPase complex. *Biosci. Biotechnol. Biochem.*, **57**, 1530–1535.