Abteilung Molekulare Botanik

(Leiter Prof. Dr. Axel Brennicke)

Universität Ulm

Characterization of mRNA processing and transcript stability in mitochondria of higher plants

Dissertation

zur Erlangung des Doktorgrades (Dr. rer. nat.)
an der Fakultät für Naturwissenschaften
der Universität Ulm

vorgelegt von

Josef Kuhn

aus Marktoberdorf

2002

Amtierender Dekan der Fakultät für Naturwissenschaften:

Prof. Dr. Wolfgang Witschel

Erstgutachter:

PD Dr. Stefan Binder, Abteilung Molekulare Botanik, Universität Ulm

Zweitgutachter:

Prof. Dr. Klaus-Dieter Spindler, Allgemeine Zoologie und Endokrinologie, Universität Ulm

Datum der Promotion:

17. Mai 2002

Die Arbeiten im Rahmen der vorgelegten Dissertation wurden in der Abteilung Molekulare Botanik der Universität Ulm durchgeführt und von Herrn PD Dr. Stefan Binder betreut.

Ulm, den 08.02.2002

Contents

1 INTRODUCTION 1
1.1 The mitochondrion 1
1.1.1 The evolution of mitochondria 1
1.1.2 The mitochondrial genome 2
1.2 Transcription of the mitochondrial genome 3
1.3 RNA processing and degradation 5
1.3.1 mRNA stability in the nuclear/cytosolic compartment of eukaryotes 5
1.3.2 The regulatory mechanisms of mRNA stability in *E. coli* 6
1.3.3 mRNA processing and degradation pathways in chloroplasts 8
1.3.4 mRNA turnover in mitochondria 11
1.4 Objective of the study 14
2 RESULTS 15
2.1 An mRNA helicase (AtSUV3) is present in *Arabidopsis thaliana* mitochondria 15
2.2 Transcript lifetime is balanced between stabilizing stem-loop structures and degradation-promoting polyadenylation in plant mitochondria 17
2.3 5' meets 3': head to tail joined mRNAs in plant mitochondria 20
2.4 RT-PCR analysis of 5' to 3'-end-ligated mRNAs identifies the extremities of *cox2* transcripts in pea mitochondria 22
3 DISCUSSION 24
3.1 AtSUV3 is a mitochondrial RNA helicase 24
3.2 Non-encoded nucleotides at the 3' end of plant mitochondrial transcripts 25
3.3 Head to tail connected mRNAs are present in the plant mitochondrial steady state RNA 26
3.4 CR-RT-PCR is an appropriate method for the simultaneous identification of 5' and 3' transcript ends and the detection of non-encoded nucleotides 29
4 SUMMARY 31
5 REFERENCES 33
6 APPENDIX 43
6.1 Presentation of the results 43
6.1.1 An mRNA helicase (AtSUV3) is present in *Arabidopsis thaliana* mitochondria 44

6.1.2 Transcript lifetime is balanced between stabilizing stem-loop structures and degradation-promoting polyadenylation in plant mitochondria 51
6.1.3 5’ meets 3’: head to tail linked mRNAs in plant mitochondria 64
6.1.4 RT-PCR analysis of 5’ to 3’-end-ligated mRNAs identifies the extremities of *cox2* transcripts in pea mitochondria 81
6.2 Curriculum vitae 90
6.3 Publications 92
6.4 Acknowledgements 93
6.5 Deutschsprachige Zusammenfassung 94

1 INTRODUCTION

1.1 The mitochondrion

1.1.1 The evolution of mitochondria

According to the generally accepted endosymbiont theory mitochondria and chloroplasts are descendants of bacteria-like organisms. This theory was first developed for the origin of chloroplasts (Schimper, 1883), and subsequently for mitochondria (Wallin, 1927). By means of molecular biological analyses this theory has been essentially verified (Margulis, 1970; Gray, 1989; Gray, 1992). Based on this theory mitochondria have a common ancestor with aerobic α-proteobacteria (purple bacteria) that seem to have been embedded in a protoeukaryotic cell, a probably anaerobic single cell organism with high similarity to nowadays archaea. The symbiosis preadapted the cells to a change from a reducing to an oxidizing atmosphere about 1.5 billion years ago. The cells were now not only able to circumvent the toxicity of oxygen by metabolizing it into harmless by-products, but also to generate energy in a process called respiration (de Duve, 1996). The origin of plastids is based on a second symbiosis event, in which a progenitor of the present cyanobacteria was integrated into a eukaryotic cell. New studies provide evidence for a single origin of chloroplasts in red and green algae, from which the green plants evolved (Moreira *et al.*, 2000; Palmer, 2000). An analogous singular origin of mitochondria is supported by several physiological and biochemical studies (Gray, 1999a; Gray, 1999b).

In recent years a new hypothesis has been proposed to explain the origin of eukaryotic cells (Martin and Müller, 1998). Based on comparing sequencing data from various genome sequencing projects, biochemical pathways and intracellular networks it is postulated that the ancestral host cell was an anaerobic, strictly autotrophic and strictly hydrogen-dependent cell. This cell with high similarity to present archaea is believed to have used geologic hydrogen for its methanogenic metabolism. With the source of geologic hydrogen steadily running dry the proto-eukaryotic cells survived by starting a symbiosis with eubacteria based on syntrophy as a driving force for the initiation of the endosymbiontic process (Doolittle, 1998; Travis, 1998). In this model it is assumed that an anaerobic progenitor of extant α-

proteobacteria excreted hydrogen and carbon dioxide as waste products of anaerobic fermentation. The archaean partner is supposed to have used the secreted hydrogen and carbon dioxide as sole sources for energy and carbon (Martin and Müller, 1998). The loss of hydrogen in the respective environment generated a selective force for a closer association between the two partners, which led to the endosymbiosis of the eubacteria like organism. With the change of the host cell from autotrophy to heterotrophy the dependence on hydrogen became decrepit. Therefore the use of increasing atmospheric oxygen in conjunction with respiratoric ATP synthesis is proposed to be an advantageous way out (Martin and Müller, 1998). Through these mechanisms the interactions became increasingly stronger and the morphology of the incorporated symbiont adapted.

1.1.2 The mitochondrial genome

In line with the above mentioned symbiosis events an active gene transfer has occurred between the different symbiont genomes. This transfer was mainly orientated towards the host cell nucleus, while some parts of the symbiont's genome were lost. In an extreme case some organelles, the so-called hydrogenosomes, have completely lost their genome (Müller, 1993). The availability of numerous completely sequenced mitochondrial genomes (chondriomes) has contributed to confirm the eubacterial origin of mitochondria. The bacterial genome sequence with the highest similarity to mitochondria, *Rickettsia provazekii* (Andersson *et al*., 1998; Gray, 1998) and the most bacteria-like mitochondrial genome sequence of *Reclinomonas americana* (Lang *et al*., 1997) illustrate the evolutionary relationship between the mitochondrial genomes and its prokaryotic relative. The mitochondrial genome of *Reclinomonas* only contains 11.6% of the genes present in the *Rickettsia* genome, with 18 of the 97 genes to be unique among investigated mitochondrial sequences.

In recent years substantiated progress has been made to elucidate the land plant mitochondrial genetic information. The completely sequenced mitochondrial genome of the Spermatophyte *Arabidopsis thaliana* contains 57 genes on 366,924 bp (Unseld *et al*., 1997). A similar amount of genes was found in the chondriome of *Beta vulgaris*, whereas the chondriome of the liverwort *Marchantia polymorpha* contains 94 genes (Oda *et al*., 1992; Kubo *et al*., 2000). In the various eukaryotes mitochondrial genome sizes are highly variable, ranging from 15 to 17 kbp for animal

genomes (Anderson *et al*., 1981; Bibb *et al*., 1981) to 200 to 2,400 kbp for the land plant genomes (Ward *et al*., 1981; Palmer and Herborn, 1987).
Although only a tiny fraction of the estimated total 2,900 mitochondrial proteins is encoded on the organellar DNA, they are essential for the function and maintenance of these organelles. Therefore the organelles have to contain a complete genetic system for the expression and inheritance of their genetic information.

1.2 Transcription of the mitochondrial genome

Transcription of the about 50 to 60 genes found in higher plant mitochondria has been studied by northern blot analyses, primer extension approaches and S1 nuclease mapping. These studies revealed monocistronic as well as polycistronic transcription units. Some of the latter are conserved between a number of plant species. One example is the *rpl5-rps14-cob* tricistronic operon, which is present in pea, rapeseed, potato and in *A. thaliana*, with the latter two species containing *rps14* pseudogenes (Aubert *et al*., 1992; Brandt *et al*., 1993; Ye *et al*., 1993; Quiñones *et al*., 1996; Hoffmann *et al*., 1999). Referring to the complete mitochondrial genome of *Arabidopsis thaliana* more than 40 genes are located on clusters. As a consequence it is highly likely that many genes of the mitochondrial genome of *Arabidopsis thaliana* are transcribed within multipartite transcription units (Dombrowski *et al*., 1998).
As a consequence of the multicistronic transcription units the mitochondrial transcripts feature very complex mRNA patterns. Surprisingly, even solitary genes with no obvious transcription partner are often represented in multiple mRNAs. *In vitro* capping analyses of plant mitochondrial mRNA revealed the activity of multiple promoters, which contributes to the complexity of the transcriptome. Striking examples for genes transcribed into multiple RNAs are the *atp9* and the *cox2* gene from maize, which are transcribed from six and five promoters respectively. As a model for promoters in dicotyledonous plant mitochondria the pea *atp9* promoter has been investigated in detail (Binder *et al*., 1995; Dombrowski *et al*., 1999). An 18 bp element ranging from –14 to +4 corresponding to the transcription initiation site (+1) was identified as a fully active promoter sequence with a highly conserved nonanucleotide motif ranging from position –7 to +2 and a well conserved motif termed AT-box from –14 to –9. The nonanucleotide motif encompasses a 5'-CRTA-3'

tetranucleotide (-7 to -4) being part of almost all higher plant mitochondrial promoters. This accounts also for monocotyledonous mitochondrial promoters, which consist of a central domain (-7 to +5) and an upstream domain centered around positions -11/-12 (Rapp and Stern, 1993; Rapp *et al*., 1993).

Transcription is most likely performed by a nuclear encoded phage-type RNA-polymerase (Hedtke *et al*., 1997). As not all mitochondrial genes or transcription units are preceded by a nonanucleotide motif-like promoter, it is speculated that an additional class of RNA-polymerase exists in plant mitochondria. In *Reclinomonas americana* subunits of an bacteria-like RNA-polymerase were identified (Lang *et al*., 1997), but till now it is unclear whether these genes are expressed *in vivo*. No homologous genes have been found in the chondriome of any other organism, suggesting that they have been lost during evolution.

With the situation at the 5' end of mRNAs becoming increasingly clearer, little is known about the termination of transcription. S1 nuclease protection experiments identified several transcript termini immediately downstream of inverted repeats, which can fold into stem-loop structures on the RNA level (Schuster *et al*., 1986). Consequently speculations arose that these structures could function as transcription terminators, similar to their function in bacteria (Lesnik *et al*., 2001). A functional analysis of the double inverted repeat of the pea *atp9* gene in an *in vitro* transcription system revealed that mitochondrial transcription proceeds through this structure without impediments and gave no indication for transcription termination (Morikami and Nakamura, 1993; Dombrowski *et al*., 1997). A consecutive analysis of several other plant mitochondrial inverted repeats corroborated that these structures are not linked to termination of transcription (Kuhn and Binder, unpublished results). Until now it remains unclear how transcription is terminated in plant mitochondria. Further investigations of the pea *atp9* inverted repeat uncovered that it acts probably as a processing signal and also contributes to the stability of the RNA (Dombrowski *et al*., 1997). Such processing events, which also take place in the 5' untranslated region of some genes further complicate the pattern of the corresponding transcripts described above (Binder *et al*., 1995).

1.3 RNA processing and degradation

The term “RNA processing” includes a variety of procedures, where the structure as well as the sequence and herewith the information content of an RNA can undergo a crucial change. In conjunction with this pre-RNAs are maturing to functional ribosomal RNAs (rRNAs), transfer-RNAs (tRNAs) or translatable messenger RNAs (mRNAs). The processes linked to these changes are the editing of specific nucleotides (Brennicke *et al*., 1999; Giegé and Brennicke, 1999), the endonucleolytic cleavage of multicistronic precursors, the splicing of introns (Michel *et al*., 1989), endo- and exonucleolytic trimming of 5’ and 3’ ends (Binder *et al*., 1995; Dombrowski *et al*., 1997; Kunzmann *et al*., 1998) and the addition of non-encoded nucleotides to the 3’ end of mRNAs (cp. 1.3.2 to 1.3.4). For many of these processes the structural basis of the respective RNA is an important feature. As an example, inverted repeat sequences in the 3’ **u**n**t**ranslated **r**egion (UTR) of many procaryotic and organellar mRNAs can form stable stem-loop structures, which seem to mediate between the stability and degradation of mRNAs (cp. 1.3.2 to 1.3.4).

1.3.1 mRNA stability in the nuclear/cytosolic compartment of eukaryotes

Similarly to prokaryotes and organelles, nuclear mRNAs undergo special processing events during and immediately after transcription. After the addition of a 7-methylguanosine to the 5’ end (capping) the nascent pre-mRNA is cleaved endonucleolytically at an A-rich sequence in the 3’ UTR prior to the addition of the poly(A) tail (Flaherty *et al*., 1997; Proudfoot, 2001). After cotranscriptional splicing of the introns the mRNA gets mature with several proteins binding to it. The **c**ap **b**inding **c**omplex (CBC) attaches to the cap structure at the 5’ end and **p**oly(**A**) **b**inding **p**roteins (PABP) get into close contact with the poly (A) tail prior to export of the ribonucleoprotein complex to the cytoplasm (Mitchell and Tollervey, 2000a). Binding of translation initiation factors, e.g. eIF4G to the CBC mediates in interaction with further factors the first round of translation of the gene. In advance of the subsequent rounds of translation the CBC is replaced by the translation initiation factor eIF4E and the PABP gets in close contact to eIF4G (Mitchell and Tollervey, 2001). This interaction results in the formation of a complex that is able to circularize the messenger RNA molecule *in vitro* (Wells *et al*., 1998). These processes are

interpreted as a proof-reading mechanism that enhances the recognition of intact and correctly processed mRNA by the translation machinery.

Transient disruption of this complex is thought to allow deandenylation at the 3' end through a **de**adenylating **n**uclease (DAN) and the following degradation of the mRNA may be performed in two alternative pathways. On the one hand the deadenylated mRNA triggers the removal of the protecting 5' cap-structure with a subsequent 5' to 3' exonucleolytic degradation of the mRNA. On the other hand a decay may occur independent of decapping. In this case a multiproteincomplex binds to the 3' end and digests the mRNA exonucleolytically from 3' to 5'. (Mitchell and Tollervey, 2000a). In the nucleus of *Saccharomyces cerevisiae* this complex, called the exosome consists of several 3' to 5' exonucleases as well as RNA-helicases (Mitchell *et al*., 1997; Mitchell and Tollervey, 2000b). Latest results indicate a functional link between the exosome and an mRNA decapping enzyme (Wang and Kiledjian, 2001). The interaction is suggested to take place after an almost 3' to 5' degradation by the exosome and the remaining cap structure is hydrolyzed by the decapping enzyme DcpS. The exosome complex is also involved in the 3' processing of the 5.8S ribosomal RNA, small nucleolar RNAs (snoRNAs) and small nuclear RNAs (snRNAs) (Allmang and Tollervey, 1998).

For some mRNAs an alternative degradation pathway has been reported, in which independent of the adenylation status, the mRNA is cleaved endonucleolytically and afterwards degraded (Beelman and Parker, 1995).

1.3.2 The regulatory mechanisms of mRNA stability in *E. coli*

Analogous to the exosome in *Saccharomyces cerevisiae* and mammals, a major component of the RNA degradation in *E. coli* is the degradosome, a multiproteincomplex. As an important part of this complex the endoribonuclease E (RNase E) interacts with all other participating proteins. These proteins are bound to the C-terminal part of RNase E, whereas the catalytically active N-terminal part mediates the localization of the entire complex to the cytoplasmic membrane (Vanzo *et al*., 1998; Liou *et al*., 2001). The enzyme cleaves the mRNA substrates upstream of stabilizing stem-loop structures, which act as thermodynamically stable secondary structures in the rho-independent termination of transcription and which inhibit 3' to 5' exoribonucleases to degrade the RNA (Birnstiel *et al*., 1985; McLaren *et al*., 1991; Hajnsdorf *et al*., 1995; Nudler *et al*., 1995; Carpousis *et al*., 1999). Surprisingly

RNase E prefers a 5' end for its activity, whereas monophosphorylated ends are preferred over triphosphorylated ends (Mackie, 1998). Circularized RNAs bearing well-characterized RNase E cleavage sites are resistant to cleavage of the enzyme *in vitro* and are about 4-6-fold more stable *in vivo* than their linear counterpart (Mackie, 2000). Moreover, RNase E has an additional function in the maturation process of 5S and 16S rRNAs (Bessarab *et al*., 1998; Li *et al*., 1999).

Polynucleotide phosphorylase (PNPase), the second major constituent of the exosome was identified to be associated with RNase E (Carpousis *et al*., 1994). It is one of the major single strand 3' to 5' exoribonucleases in mRNA decay. This phosphorolytic enzyme uses inorganic phosphate to remove nucleotides from the 3' end, yielding nucleoside 5' diphosphates (Deutscher and Reuven, 1991). The third component of the degradosome is an ATP dependent DEAD-box RNA-helicase, RhlB, which is responsible for the unwinding of internally structured regions that force enzymes such as the PNPase to pause during degradation (McLaren *et al*., 1991; Py *et al*., 1996). The function of enolase, a glycolytic enzyme and a further component of the degradosome are unclear up to now. That holds also true for the additional substoichiometrical components of the complex, the chaperons DnaK and GroEL and the polyphosphate kinase (PPK). The latter is thought to maintain an appropriate microenvironment on the basis of ATP and inorganic phosphate for the degradosome (Miczak *et al*., 1996; Blum *et al*., 1997). An mRNA degrading complex with the components described above has also been isolated from *Rhodobacter capsulatus*, an α purple bacterium (Jäger *et al*., 2001).

Beside the degradosome other 3' to 5' exoribonucleases play major roles in mRNA degradation: RNase II accounts for about 90% of the exonucleolytic activity in *E. coli* and the oligoribonuclease, which is required to break down the remaining small 2-5 nt long mRNA fragments (Ghosh and Deutscher, 1999; Steege, 2000).

Polyadenylation has been believed for a long time to be associated exclusively with eukaryotic mRNAs, where extensions between 80 and 200 adenosines can be observed (Steege, 2000). Nevertheless, poly (A) tails have also been detected in bacteria (Hajnsdorf *et al*., 1995; O'Hara *et al*., 1995; Xu *et al*., 1995; Li *et al*., 1998). The polyadenylated RNA with a length between 10 and 50 nt accounts for only a very small fraction of the RNA population in the cell. This fraction increased several fold in mutant bacterial cells lacking exoribonuclease activities. Moreover, mRNA in cells lacking RNA polyadenylation have a dramatic increase in half-life time (Cohen,

1995). These results suggest that in contrast to the nuclear and cytoplasmic compartment of the eukaryotic cells, the addition of poly(A) tails to the 3' end of bacterial mRNAs promotes their degradation. That applies also for mRNAs whose 3' ends are sequestered in secondary structures (Blum *et al*., 1999; Carpousis *et al*., 1999). The poly(A) tail is supposed to provide a toehold for PNPase and RNase II which then bind to the 3' end of the RNA and efficiently initiate degradation. It has been suggested, that repeated polyadenylation and exonucleolytic attack could be required to overcome the secondary structures of certain intermediates (Coburn and Mackie, 1998). It should be noted, that polyadenylation is not restricted to mRNAs, but has also been observed at precursor and mature forms of rRNAs and small RNAs, although the functions of these adenosine extensions are unclear at present (Li *et al*., 1998; Mohanty and Kushner, 2000a). Beside homopolymeric tails consisting of adenosines also heteropolymeric tails with C, U and G residues occasionally occur in some stages of the cell cycle (Cao *et al*., 1996).

The enzyme responsible for the posttranscriptional addition of poly(A) tails is the poly(A) polymerase I (PAPI), which accounts for about 90 to 95% of the poly(A) tails in wild type *E. coli* (August *et al*., 1962; Cao and Sarkar, 1992; O'Hara *et al*., 1995; Mohanty and Kushner, 1999). Recently poly(A) polymerase activity was also reported to the PNPase, which seems to add non-encoded nucleotides to the 3' end of various RNAs (Mohanty and Kushner, 2000b).

Analogous to the interaction of poly(A) binding proteins with other proteins of the nuclear/cytosolic ribonucleoprotein complex, such an interaction has been shown for poly(A) binding proteins and components of the degradosome in *E. coli* (Feng *et al*., 2001). One class of these proteins, CspE, impede the exonucleolytic decay of PNPase and inhibit internal cleavage through RNase E. The second class, the ribosomal protein S1, seems to link the functions between degradation and translation in prokaryotic cells. This theory is further supported by the finding, that polyadenylation of mRNAs does not effect protein synthesis *in vitro* (Lee and Cohen, 2001), raising the possibility that poly (A) tails, while accelerating the decay of mRNAs, also lead to a compensatory increase in translation efficiency.

1.3.3 mRNA processing and degradation pathways in chloroplasts

Most plastid RNAs undergo extensive processing, since many of them are part of polycistronic transcription units, which have to be converted into monocistronic

transcripts to increase translation efficiency and to allow an individual regulation of transcript stability (Barkan, 1988; Westhoff and Herrmann, 1988; Kim *et al*., 1993).

For many plastid mRNA molecules a processing step at the 5' end seems to be an important step. For example the barley *rbcL* mRNA has to be cleaved endonucleolytically in the 5' UTR to associate with polysomes (Reinbothe *et al*., 1993). The same applies for the *rbcL*, *psbA*, *psbB*, *psbD* and the *petD* transcripts in *Chlamydomonas reinhardtii* chloroplasts (Sakamoto *et al*., 1994; Shapira *et al*., 1997; Bruick and Mayfield, 1998; Nickelsen *et al*., 1999; Vaistij *et al*., 2000). Concerning the *petD* transcript, a 5' to 3' exonucleolytic activity has been described to generate a mature transcript (Drager *et al*., 1998; Drager *et al*., 1999). Moreover several analyses suggest that almost all mRNAs are processed at the 5' end in *Chlamydomonas reinhardtii*, since *in vitro* capping is not possible due to the lack of a triphosphate at the 5' termini (Monde *et al*., 2000). Up to now no specific enzyme responsible for the 5' processing has been identified, whereas several proteins with 5' UTR binding activity were found (Alexander *et al*., 1998; Monde *et al*., 2000). These proteins are thought to mediate stabilizing and translation activating influences on the respective mRNAs.

The 3' processing of plastid mRNAs plays an important role for the realization of the genetic information. Like in *E. coli*, most plastid mRNAs have inverted repeats in their 3' UTR, which are supposed to form stem-loop structures. In contrast to bacteria these structures are not involved in transcription termination *in vitro* and *in vivo* (Stern and Gruissem, 1987; Stern *et al*., 1989; Rott *et al*., 1996). Although most plastid mRNAs have homogeneous 3' ends about 3 to 10 nt downstream of the stem-loops (Stern and Gruissem, 1987; Hayes *et al*., 1996; Rott *et al*., 1998a). The removal of excessive nucleotides resulting from transcription occurs in a two-step reaction. After an endonucleolytic cleavage about 10 to 20 nt downstream of the stem-loop the remaining nucleotides are removed in a 3' to 5' exonucleolytic reaction (Stern and Kindle, 1993; Hayes *et al*., 1996). This maturation step has to be fulfilled quite thoroughly, because the lack of the hairpin results in fast degradation of the transcript and only mRNAs with a correct 3' end are translated further on (Drager *et al*., 1996; Rott *et al*., 1998b).

Apart from several proteins like RNP 28, RNP 33 or RNP 55 binding to the 3' UTR and the stem-loop structure, a high molecular weight complex (HMW) could be enriched from spinach chloroplasts, which seems to be responsible for the correct 3'

processing (Stern *et al*., 1989; Chen and Stern, 1991a; Schuster and Gruissem, 1991; Chen *et al*., 1995; Hayes *et al*., 1996). The complex contains a protein with a molecular mass of 100 kDa, which was identified as a 3' to 5' exonuclease with high similarity to the PNPase from *E. coli* (Hayes *et al*., 1996). The complex with a molecular mass of 580 to 600 kDa has recently been shown to consist of a homo-multimer of several PNPase proteins, which is quite distinct from the *E. coli* degradosome (Baginsky *et al*., 2001). In this approach no other chloroplast proteins were found to associate with chloroplast PNPase during affinity chromatography. Moreover, a protein with a molecular mass of 67 kDa, which was suggested to correspond to the RNase E from *E. coli*, was identified as the α and β subunit of the chloroplast chaperon GroEL (Hayes *et al*., 1996; Baginsky *et al*., 2001). To date, the endonucleolytic activity described by Hayes *et al*. (1996) has not been substantiated on the protein level. Furthermore, a database analysis of proteins homologous to *E. coli* RNase E revealed that the respective proteins lack the C-terminal domain of the *E. coli* protein that serves as a platform for the assembly of the degradosome in *E. coli* (Vanzo *et al*., 1998; Baginsky *et al*., 2001). Taken together, RNA processing and degradation differ in several aspects between chloroplasts and *E. coli*.

Despite the differences, similarities were observed in terms of polyadenylation. In chloroplasts non-encoded poly (A) tails of about 45 nt length were initially demonstrated in maize (Haff and Bogorad, 1976). The positions of the polyadenylation were later determined to be mainly upstream of the 3' inverted repeat and at the mature 3' end (Kudla *et al*., 1996; Lisitsky *et al*., 1996). Unclear is at present the length and composition of the extensions. Detailed analyses have only been performed in *Chlamydomonas* chloroplast, where almost homopolymeric (>98 % adenine) extensions between 5 and 85 nt in length were found (Tomine *et al*., 2000). Poly (A) extensions were not only observed at mRNAs, but were also found at tRNAs and rRNAs (Tomine *et al*., 2000). Like in *E. coli* polyadenylation in chloroplasts seems to destabilize RNA molecules at least *in vitro* (Kudla *et al*., 1996; Lisitsky *et al*., 1997a). This is consistent with the finding that the blocking of polyadenylation inhibits mRNA degradation (Lisitsky *et al*., 1997b). It is remarkably that PNPase has a much higher affinity for polyadenylated versus non-polyadenylated RNAs (Lisitsky *et al*., 1997a).

Concerning a polyadenylating activity in chloroplasts, PNPase is discussed as the only activity being able to polyadenylate mRNA molecules, since the polyadenylating

activity could not be separated from the PNPase activity during purification (Yehudai-Resheff *et al*., 2001). This is, however completely contradictory to former results of the same group, where a separation of a polyadenylating activity and the PNPase was reported (Lisitsky *et al*., 1996). The existence of a poly(A) polymerase (PAP) like enzyme is not yet verified, although such an enzyme is postulated (Lisitsky *et al*., 1997a).

The degradation of chloroplast mRNAs is like in *E. coli* supposed to be initiated with endonucleolytic cleavages of RNAs, what is substantiated with the enrichment of intermediates during blocking of exonuclease activities and the positions of the poly(A) tails added to free 3' ends (Klaff, 1995; Kudla *et al*., 1996; Lisitsky *et al*., 1996). Several enzymes were postulated and characterized, which are involved in the endonucleolytic cleavage of transcripts (Chen and Stern, 1991b; Kudla *et al*., 1996; Yang and Stern, 1997). The fragments of the endonucleolytic digestion are polyadenylated and supposed to be degraded rapidly through 3' to 5' exonucleases, whereas the inverted repeat structure is supposed to be degraded in a 5' to 3' exonucleolytic direction (Monde *et al*., 2000).

1.3.4 mRNA turnover in mitochondria

The steady state levels of plant mitochondrial mRNAs are most likely determined by a combination of transcriptional rates and the stability of the transcripts (Giegé *et al*., 2000). The latter is influenced by processes involving RNA secondary structure elements, which function as processing signals and stability elements. As in bacteria and in chloroplasts, these elements consist of inverted repeats able to form stem-loop structures. Such inverted repeats have been identified in the 3' non-translated regions of several plant mitochondrial genes, and the localization of 3' mRNA termini immediately downstream of these structures raised the possibility that they function analogous to those in bacteria and chloroplasts (Dewey *et al*., 1986; Schuster *et al*., 1986; Morikami *et al*., 1987; Bonen and Bird, 1988; Wahleitner and Wolstenholme, 1988; Gualberto *et al*., 1990; Liu *et al*., 1992; Dell'Orto *et al*., 1993). The presence or absence of such stem-loop structures in otherwise almost identical *cob* transcripts correlates with differing steady state abundances in certain rice and wheat lines, implicating a function of these stem-loops in transcript stabilization (Kaleikau *et al*., 1992; Saalaoui *et al*., 1990).

A detailed functional analysis of the pea *atp9* inverted repeat revealed no function of the double stem-loop as transcription terminator at least *in vitro* (Dombrowski *et al*., 1997). Analogous results were obtained in studies with respective structures located in the 3' UTR of the *Oenothera atp1* gene, the *Oryza cob-I* gene and the *potato atp9* gene (Kuhn and Binder, unpublished results). *In vitro* processing assays of the pea *atp9* double stem-loop suggested its involvement as a processing signal and stability element, since the 3' ends of the processing products coincide with those mapped *in vivo* (Dombrowski *et al*., 1997).

Beside these stabilizing processes there are also degradation-promoting mechanisms operative in plant mitochondria. Recently the degradation of CMS-related *atp1-orf522* transcripts in sunflower was shown to be correlated with preferential polyadenylation of this (Gagliardi and Leaver, 1999). Shortly later polyadenylation of *cox2* transcripts was reported in maize mitochondria, where non-encoded adenosines were found to be added at multiple sites throughout the transcript, with 2 major sites coinciding with mapped 3' ends (Lupold *et al*., 1999). In this case the destabilizing effect of the polyadenylation was very weak *in vitro*. In turn, a relatively strong stimulating effect of polyadenylation on *atp9* mRNA degradation was observed in potato mitochondria, where a 3' to 5' exonucleolytic activity is not significantly impeded by secondary structures present in RNA substrates (Gagliardi *et al*., 2001).

In contrast to all other genetic systems, where several components of the processing complex have been identified and characterized, virtually nothing is known about the constituents of the RNA processing machinery in plant mitochondria. In yeast mitochondria, an RNA helicase called SUV3 is part of a degradosome-like complex, that includes also an exoribonuclease and that may control mRNA stability in these organelles (Stepien *et al*., 1992; Anderson and Parker, 1996; Margossian *et al*., 1996; Carpousis *et al*., 1999). The interaction of such a putative multi-protein complex with an AU-rich dodecamer motif in the 3' UTR determines whether a transcript gets degraded or stabilized. The 3' ends of all stable and processed mRNAs is mapped exactly at the dodecamer motif implicating an accurate processing step with an endonucleolytic removal of the downstream sequence being very likely (Min and Zassenhaus, 1993; Margossian and Butow, 1996). Similar to above mentioned genetic organisms short oligo(A) tails were detected at the 3' end of RNAs in yeast mitochondria, being about 8 nt long and predominantly located at

the large rRNAs (Yuckenberg and Phillips, 1982). However, the appearance and function of such proteins and RNA modifications in plant mitochondria remains to be determined, since RNA processing in mitochondria of plants and yeast seems to be different.

1.4 Objective of the study

In this thesis polyadenylation of several plant mitochondrial mRNAs should be analyzed in different species of dicotyledonous plants. A possible function of the polyadenylation should be investigated in *in vitro* processing assays. To gain an insight into the enzymatic activities involved in the addition of these extensions, it was aimed to identify the lengths and composition of mRNA 3' ends. In addition the analysis of the several transcript extremities by alternative methods should give general informations about the structure and formation of 3' ends and the addition of non-encoded nucleotides.

Beside the structure of the RNAs, proteins potentially involved in plant mitochondrial 3' processing and RNA stability should be identified and studied.

2 RESULTS

In the present thesis the analysis of plant mitochondrial processing events at the 3' and the 5' terminus is followed in four approaches. After the identification and characterization of the putative RNA-helicase AtSUV3 in *Arabidopsis thaliana* the occurrence and the role of polyadenylated mRNAs on transcript stability was investigated for the *atp9* gene from *Pisum sativum* and the *atp1* gene from *Oenothera berteriana*. In this context the composition and the lengths of non-encoded extensions could be elicited in pea mitochondrial *atp9* transcripts *in vivo*. During these analyses a covalent linkage between 5' and 3' transcript termini of the two genes was detected. On the basis of this finding a new method for the simultaneous mapping of 5' and 3' ends and the detection of non-encoded nucleotides at the 3' end of the respective transcript was established.

2.1 An mRNA helicase (AtSUV3) is present in *Arabidopsis thaliana* mitochondria

Since the proteins involved in mitochondrial mRNA processing and degradation have yet to be identified, an RNA helicase was identified on the basis of significant similarity to an already characterized protein from yeast mitochondria (SUV3) (cp. 1.3.4; Gagliardi *et al*., 1999, reprint see 6.1.1). A *suv3* homologue sequence was found on chromosome VI of *Arabidopsis thaliana*, sharing 39% and 45% identical amino acids with the yeast and human homologues, respectively (Altschul *et al*., 1997; Gagliardi *et al*., 1999, Fig.1). Therefore the identified gene was termed *atsuv3*. Since no EST sequence corresponding to the predicted *atsuv3* gene was available, a cDNA was amplified with primers at the respective ends of the predicted coding region resulting in a reading frame of 1290 bp. Sequencing of the amplified product confirmed the computer predictions, with only the first intron extending 120 bp further downstream and the 5' end of the fourth intron being situated 50 bp further downstream. As a large part of the C-terminus was missing in AtSUV3 in comparison to SUV3 sequences of other organisms a 3' RACE analysis was performed. This identified seven additional exons eliminating the predicted stop codon and extending the orf to 1713 bp with a calculated molecular mass of 63.6 kDa. To verify the predictions and cDNA sequencing a stringent Southern blot analysis was performed,

in which the single genomic location was confirmed with the fragments of the digested DNA matching exactly the computer predictions (Gagliardi *et al*., 1999, Fig. 2). Although the AtSUV3 protein is between 76 and 215 amino acids shorter than the homologues from yeast, human or nematode, all domains of a typical DExH RNA helicase are present in the central part, which is well conserved between all mentioned organisms. To study the expression of the *atsuv3* gene Northern blot analyses were carried out with 10 µg total RNA and 5,7 µg mRNA from *A. thaliana* total plants. Only a faint signal corresponding to an *atsuv3* transcript with a size of 2.3-2.5 kbp could be detected in the fraction of poly(A)$^+$ RNA, indicating a very weak expression of this gene (Gagliardi *et al*., 1999, Fig. 3a). In order to confirm the transcript size resulting from Northern blot and 3' RACE analyses primer extension experiments were performed initiated from an oligonucleotide annealing 40 bp downstream of the initiation codon. A major product of about 540 nt upstream of the initiation codon is consistent with a total size of the transcript of 2.5 kbp (Gagliardi *et al*., 1999, Fig 3b). All experiments described were done by myself with basic assistance of Ulrike Spadinger.

Dominique Gagliardi undertook additional experiments to further characterize the protein. To proof the mitochondrial localization of the AtSUV3 protein, he performed a cell fractionation experiment with transgenic plants expressing a fusion protein consisting of the AtSUV3 N-terminal region and the reading frame of the glucuronidase reporter gene. Distribution of the GUS protein and its activity with the parallel detection of mitochondrial marker proteins in the discrete cell fractions clearly localizes the AtSUV3 protein in mitochondria of *Arabidopsis thaliana* and *Solanum tuberosum* (Gagliardi *et al*., 1999, Fig.4). This localization has been verified through a transient protoplast transformation with a construct consisting of a fusion of the 5' part of the *atsuv3* cDNA and the reading frame for the expression of a green fluorescence protein (GFP) (Kuhn and Binder, unpublished results). The conservation of all major sequence elements of DExH RNA helicases, including an ATP-binding domain (amino acids 96 to 103), suggests a function of the AtSUV3 similar to the SUV3 in yeast. After overexpression of AtSUV3 in *E. coli* the enzyme showed an ATP hydrolysis activity in dependence of mitochondrial RNA clearly indicating an ATPase characteristic for RNA helicase (Gagliardi *et al*., 1999, Fig 5).

2.2 Transcript lifetime is balanced between stabilizing stem-loop structures and degradation-promoting polyadenylation in plant mitochondria

To elucidate the mechanisms of RNA processing in plant mitochondria, polyadenylation and its effects on transcript stabilities was investigated in *Pisum sativum* and *Oenothera berteriana*.

The pea *atp9* gene is transcribed into three different mRNA species with different 5' termini but identical 3' ends. The latter have been mapped in two independent nuclease S1 protection experiments within or immediately downstream of the second of two consecutive inverted repeats (Morikami and Nakamura, 1993; Dombrowski *et al*., 1997; cp. 1.3.4). To investigate whether non-encoded poly(A) tails are present at these transcript termini, 3' RACE analyses including two rounds of PCR with an oligo d(T) adapter primer and two nested *atp9* specific primers were performed with total pea mitochondrial RNA (Kuhn *et al*., 2001, Fig 1; reprint see 6.1.2). In 18 clones non-encoded poly(A) tails were identified exactly at previously mapped 3' ends. All cDNAs contain a C to U exchange at position +20 relative to the translation start codon (+1) originating from an RNA editing event. In four clones non-encoded cytidine, guanosine and thymidine nucleotides were found. The 3' RACE analysis was done with the assistance of Ulrike Tengler.

A similar result was obtained in an analogous analysis of *Oenothera atp1* transcripts. In eleven cDNA clones the addition of poly(A) tails was found mainly downstream of the second inverted repeat, corresponding exactly to previously mapped 3' termini (Schuster *et al*., 1986; Kuhn *et al*., 2001, Fig.2).

In recent studies it was suggested, that transcripts with poly(A) tails decrease mRNA stability in plant mitochondria (Gagliardi and Leaver, 1999; Lupold *et al*., 1999). This effect would counteract the function of the double stem-loop, which was found to act as a stability element (Dombrowski *et al*., 1997). To test the effect of polyadenylation on the stability of transcripts containing an inverted repeat structure, synthetic transcripts of the pea *atp9* 3' UTR were incubated in a pea mitochondrial *in vitro* processing system (Dombrowski *et al*., 1997). One transcript had an extension of 20 adenosines immediately downstream of the stem-loop, whereas the other transcript had a tail of the same length corresponding to genomically encoded nucleotides. Products were separated with polyacrylamid gel electrophoreses. In contrast to non-polyadenylated RNA, where a single product with a size of about 70 nt can be

observed after 60 minutes, an effective removal of the poly(A) extensions is seen already after 10 minutes. In contrast to the non-polyadenylated transcript, processing of the polyadenylated transcript results in two products of different sizes, indicating the occurrence of a processing intermediate (Kuhn *et al*., 2001, Fig.3). A similar processing pattern was observed in an identical approach with *Oenothera atp1* transcripts (data not shown).

The precise 3' ends of the *atp9 in vitro* processing products were determined by an alternative 3' RACE approach. Here anchor primers were directly ligated to the 3' end of the processed RNAs, followed by an RT-PCR analysis initiated with an oligonucleotide complementary to the anchor primer (Kuhn *et al*., 2001, Fig.4). This experiment revealed, that the 3' ends generated at the non-polyadenylated precursor are scattered over a broad range. The majority of the 3' termini are located in the previously mapped region of the *atp9* gene. The 3' ends of the smaller products resulting from the polyadenylated precursor coincide with the ends seen without poly(A) tail. The larger products generated from the polyadenylated precursor are degradation intermediates with 3' ends downstream of the second stem-loop. About 50 % of these clones contain one non-encoded adenosine, indicating that the poly(A) tail is not completely removed. The products observed after 60 minutes are slightly shorter. Their 3' ends are located within the second stem of the inverted repeat, indicating a further degradation upon longer incubation in the processing lysate.

To investigate the temporal dynamics of the *in vitro* processing and to elucidate the possibility, whether the intermediate processing product is generated endonucleolytically, processing of polyadenylated precursor was followed over a period of 90 minutes (Kuhn *et al*., 2001, Fig.5). Already after one minute a diffuse smear indicates an exonucleolytic degradation of the precursor. From 2.5 minutes on, the processing intermediate described above can be observed with the highest abundance after 10 minutes. Further incubation results in a shortening and disappearance of the intermediate, coming along with a persistent increase of the smaller processing product with a maximum abundance after 60 minutes. The observed progression strongly suggest an exonuclease to remove the downstream part of the precursor with a stalling not only in the inverted repeat, but also at the transition from the poly(A) stretch to the genomic sequence.

To investigate the stabilizing function of the stem-loop structure during the formation of the different processing products, polyadenylated and non-polyadenylated RNAs

without a stem-loop were tested in the *in vitro* processing system (Kuhn *et al*., 2001, Fig.6). The substrate without stem-loop and poly(A) tract is consecutively degraded without formation of any stable intermediate, indicating that the stem-loop is indeed responsible for the generation of stabilized products. In contrast, the presence of the poly(A) tail consisting of 21 adenosines substantially accelerates the complete turnover of the precursor RNA. Again a weak intermediate results from a stalling at the transition from the poly(A) tract to the "normal" sequence.

To further pursue the assumption that an exonuclease activity is responsible for the 3' to 5' degradation, the non-polyadenylated *atp9* precursor was labelled either at the 5' or the 3' end and the substrates incubated in the *in vitro* processing system. While a product of the expected size is seen with the 5' labelled substrate, the 3' labelled products disappear without formation of detectable processing products implicating that indeed a 3' to 5' exoribonuclease takes part in the 3' processing event.

In order to analyze the identity of the non-encoded nucleotides and the length of the extensions *in vivo*, anchor primers were ligated directly to the 3' ends of total steady-state mitochondrial RNA from pea. The products of the following RT-PCR initiated at an oligonucleotide complementary to the anchor primer were cloned and analyzed (Kuhn *et al*., 2001, Fig.7 & 8). From 39 clones sequenced, 13 contain 3' non-encoded nucleotides at sites mapped previously as 3' termini. Extensions no longer than three nucleotides were observed, consisting of adenosines but also of other nucleotides, predominantly cytidines. RNA fragments deriving from 18S rRNA or the 5' UTR of the *atp9* transcript seem to be most likely coincidentally ligated to the 3' ends.

Since only short extensions up to three nucleotides are found in the alternative 3' RACE approach, the question arose whether these extensions are indeed so short and fully functional or whether they are the result of the processing reaction. Therefore substrates without inverted repeat containing 0, 3, 10 and 21 adenosines were incubated in the pea mitochondrial lysate. While both 10 and 21 adenosines significantly accelerate degradation, no effect is observed with shorter poly(A) tracts (Kuhn *et al*., 2001, Fig.8). Analogous results were obtained with substrates containing inverted repeats. These results suggest that the short extensions detected *in vivo* are the products of processing and/or degradation.

2.3 5' meets 3': head to tail joined mRNAs in plant mitochondria

While investigating the *in vivo* 3' termini of the pea *atp9* mRNAs by ligating an anchor primer directly to the 3' end, two cDNA clones were detected, deriving from transcripts where the 5' moiety of a pea *atp9* mRNA is connected to the 3' part of the same transcript (Kuhn *et al*., 2001; cp. 2.2). Since these clones were obtained in independent ligation and amplification reactions and even from different RNA preparations, it seemed highly unlikely that the extremities of the same transcript are incidentally connected by T4 RNA ligase used in the experiment. To investigate whether such head to tail joined pea *atp9* transcripts exist *in vivo*, an RT-PCR analysis was carried out. The cDNA synthesis was initiated from 5' sequences followed by an amplification reaction across the potential 5' to 3' linkage (Kuhn and Binder, manuscript, Fig.1; cp. 6.1.3). The first amplification generated products of various sizes. Since a product of about 390 bp is expected upon linkage of the previously mapped 5' and 3' ends PCR products with sizes between 300 and 500 bp were recovered from the gel and used as template in two different additional amplification reactions. The products of the nested PCR approach were cloned and sequenced.

The analysis of 29 clones indeed revealed the *in vivo* existence of 5' to 3' end joined pea *atp9* mRNAs (Kuhn and Binder, manuscript, Fig.2). In 15 clones the 5' terminus is found at position -136 relative to the start codon. In five clones the 5' end is located at position -135, in 3 clones at position -118 and the remaining 6 clones scatter from -134 to -12. In 22 clones the 5' part is joined to 3' termini located directly downstream of the second inverted repeat or within the downstream part of the second repeat as mapped previously (Kuhn *et al*., 2001). The 3' ends of the residual clones are found in the downstream loop, a single clone contains a 3' terminus in the first stem.

Beside head to tail conformation of the mRNAs, the sequences of 16 clones revealed the presence of non-encoded nucleotides, mostly two to three adenosines, while a cytidine or guanosine are only found in single clones. These data are consistent with results found in previous studies (Kuhn *et al*., 2001). Since also a substantial amount of clones were found without non-encoded nucleotides their presence seems to be independent from the formation of head to tail joined transcripts.

To investigate these findings on a wider basis the analysis was extended to *atp1* transcripts from *Oenothera*. An analogous strategy identified head to tail *atp1* mRNA in mitochondria of this plant species (Kuhn and Binder, manuscript, Fig.4). From 32

clones sequenced only a single cDNA contained a 5' moiety starting with the transcription initiation site (-209). All other clones contain 5' ends starting at positions -134 to -131 indicating a new 5' end. The termini of the 5' parts of the four remaining clones coincide with nucleotides -199, -72 and -7. The 3' ends detected in this approach can be divided into two categories. They are either located directly downstream of the double stem-loop or in the first loop of the upstream stem-loop. The former ends, which were found in 22 clones are consistent with previously mapped 3' ends, while the latter ends have not been described before (Kuhn *et al*., 2001). None of the 3' ends in the first loop of the upstream stem has additions of non-encoded nucleotides. In contrast, 16 out of 22 clones deriving from 3' ends immediately downstream of the stem-loop contain mostly between four and five non-encoded nucleotides, in the majority adenosines.

Similar results were obtained in a control reaction with total mitochondrial RNA that have been self-ligated with T4 RNA ligase. (Kuhn and Binder, manuscript, Fig.4c) Here a more homogenous population of 5' to 3' borders could be observed. Only a single clone does not match -134 to -131 5' ends and almost all 3' ends terminate directly downstream of the double stem-loop with only a single clone ending within the downstream part of the second stem. In contrast, the length of non-encoded adenosine stretches found in these clones is more variable, ranging from one to five nucleotides.

To confirm the new 5' termini identified at positions -134 to -131, an independent experimental approach was performed. In a primer extension analysis two strong signals corresponding to products of about 230 and 300 nt perfectly correlate with the 5' ends at the transcription initiation site and the newly identified terminus. In summary, the presence of head to tail connected transcripts is not restricted to a single gene in a single plant species, but rather seems to be a more general phenomenon in plant mitochondria. All experiments in this analysis were done by myself.

2.4 RT-PCR analysis of 5' to 3'-end-ligated mRNAs identifies the extremities of *cox2* transcripts in pea mitochondria

In order to analyze the extremities of mitochondrial transcripts without stem-loop structures in the 3' UTR, an RT-PCR of in vitro circularized mitochondrial RNA was performed. An ideal candidate was the pea *cox2* gene, which is transcribed into a single transcript that has been mapped before (Moon *et al*., 1985; Kuhn and Binder, 2002). After self-ligation of pea mitochondrial RNA with T4 RNA ligase specific first-strand cDNA synthesis was initiated followed by a PCR, cloning and sequencing of the cDNA clones (Kuhn and Binder, 2002, Fig.1). This **c**ircular **R**NA-RT-PCR (CR-RT-PCR) analysis revealed 5' ends scattering over a range of 31 nt mostly downstream of those mapped previously (Kuhn and Binder, 2002, Fig.2). In contrast the 3' terminal sequences do not match the sequence formerly published. Instead it extends 44 nt further, followed by 61-64 nt, depending on the cDNA clone, not found in the published genomic sequence.

Since the 5' termini detected in this approach do not match the previously mapped ends these were re-examined in a primer extension analysis (Kuhn and Binder, 2002, Fig.3). These experiments revealed three major 5' ends in repeated experiments in the initially described region, but all differing from the previously mapped sites. They also differ from 5' ends detected in the RT-PCR approach with the circularized RNA. In order to investigate whether the identified ends indeed derive from transcription initiation, two different DNA templates were incubated with a transcriptionally active mitochondrial protein lysate (Kuhn and Binder, 2002, Fig.4). All three 5' ends seem to be transcription start sites, even though the surrounding sequences of the most upstream initiation site does not resemble to any known promoter sequence from dicot plant mitochondria (Binder *et al*., 1995; Dombrowski *et al*., 1999).

To this point the situation at the 3' end was still unclear. Although an addition of 3' non-encoded nucleotides has been reported, an attachment of a >60 nt large, non-homopolymeric sequence matching exactly to the genomic sequence is highly unlikely (Kuhn *et al*., 2001). Therefore the 3' region of the *cox2* gene was examined on the level of genomic DNA. This revealed the newly identified sequence being present in the genomic DNA (Kuhn and Binder, 2002, Fig.5; DDBJ/EMBL/GenBank accession no. AJ414385). To confirm this again on the transcript level by an independent experimental approach, an S1 protection analysis covering the 3' UTR was performed (Kuhn and Binder, 2002, Fig.6). The protected fragment of ~100 nt is

consistent with the 3' transcript termini detected in the RT-PCR analysis (see above). Surprisingly, comparison of the cDNA sequences derived from RT-PCR analysis with genomic sequences revealed, in contrast to the results presented for the pea *atp9* and the *Oenothera atp1* transcripts, no non-encoded nucleotides at the 3' ends of the *cox2* mRNA.

3 DISCUSSION

3.1 AtSUV3 is a mitochondrial RNA helicase

The steadily increasing amount of sequencing data in plants generates a cumulative number of sequences with similarity to genes in other organisms (The *Arabidopsis* Genome Initiative, 2000). Although the sequence similarities suggest identical functions, these may have changed and similarity alone may be deceptive. Thus a detailed analysis of the subcellular localisation and of biochemical properties of the gene products is obligatory. Accordingly, the complete cDNA sequence of a plant homologue of the yeast SUV3 RNA helicase, AtSUV3 has been determined and the expression of the gene was analyzed (Gagliardi *et al*., 1999; cp. 6.1.1). The prediction of a mitochondrial localisation program PSORT has been confirmed experimentally by cofractionation of a AtSUV3:GUS fusion protein with mitochondrial marker proteins and the mitochondrial localization of a AtSUV3:GFP fusion protein in mitochondria of tobacco protoplasts (Gagliardi *et al*., 1999, Fig.4; Kuhn and Binder, data not shown).

The presence of homologues *suv3* sequences in eukaryotic representatives indicates a general function of this protein in the mitochondrial RNA metabolism. In yeast the SUV3 protein is involved in 3' RNA processing and degradation of group I introns (Conrad-Webb *et al*., 1990; Margossian *et al*., 1996). Since mitochondrial genes of *Arabidopsis thaliana* as well as mammals do not contain group I introns and RNA 3' processing signals and mechanisms differ from those in yeast, the suggestion arises that the SUV3 proteins may have a more general function in mitochondrial RNA metabolism (Min and Zassenhaus, 1993; Unseld *et al*., 1997). Despite the evolutionary distance between fungal and plant enzymes the enzymatic activity of the RNA helicase probably has not altered. The plant enzyme indeed hydrolyzes ATP in an RNA stimulated manner, what is a common characteristic of ATP-dependent RNA helicases (Gagliardi *et al*., 1999, Fig.5). A further characterization of the AtSUV3 helicase *in vitro* and *in vivo*, its possible association with other proteins as well as the identification of potential endogenous RNA substrates have to be investigated to elucidate the exact function of this enzyme in mitochondrial RNA metabolism.

3.2 Non-encoded nucleotides at the 3' end of plant mitochondrial transcripts

Two independent approaches were used to identify non-encoded nucleotides at the 3' end of plant mitochondrial transcripts (Kuhn *et al*., 2001). A conventional 3' RACE analysis of pea *atp9* transcripts as well as Oenothera *atp1* transcripts predominantly detects poly(A) tracts within the downstream part of the second stem or immediately adjacent. Besides the adenosines expected from the use of an oligo(dT)-adapter primer, cytidines, thymidines and guanosines are also found in the cDNA clones, mostly found directly at or near the 3' most nucleotides. The alternative direct anchor primer ligation RT-PCR approach confirms the presence of adenosines, cytidines and uridines are found in more than half of the clones sequenced, while no non-encoded guanosines are seen in this approach. The possibility that the 3' non-encoded nucleotides might be artificially added during these experimental procedures is unlikely, since the analogous investigation of *in vitro* processing products does not reveal any nucleotide that could derive from methodological artifacts. The majority of the cDNA clones resulting from the direct anchor primer ligation to total mitochondrial RNA contains a single non-encoded nucleotide, and only few clones with maximally three posttranscriptionally added nucleotides implicate a low frequency of longer extension. *In vitro* processing assays with RNAs containing three 3' terminal adenosines demonstrated no degradation promoting effect and thus suggest the RNAs with such short extensions to be relatively stable degradation or processing products.

During the course of this thesis, polyadenylation was reported in mitochondria of various plant species. In sunflower a dicistronic mRNA associated with cytoplasmic male sterility was suggested to be specifically destabilized by polyadenylation in fertility-restored lines (Gagliardi and Leaver, 1999). In maize, polyadenylation was detected at multiple sites within *cox2* transcripts (Lupold *et al*., 1999). The degradation promoting effect of polyadenylation in maize mitochondrial extracts, was however very weak. An alternative approach has also confirmed the addition of non-encoded nucleotides to 3' transcript ends of several mitochondrial genes in maize (Williams *et al*., 2000). Recently potato mitochondrial *atp9* mRNAs were reported to have polyadenylate extensions at their 3' termini (Gagliardi *et al*., 2001).

In general a degradation promoting effect was suggested for polyadenylation of mitochondrial transcripts in these reports. This is consistent with my results obtained

with pea *atp9* RNA without a stabilizing stem-loop (Kuhn *et al*., 2001). In contrast, only an accelerated processing is observed in assays with polyadenylated substrates with stem-loop, while total degradation is only marginally increased. Thus the stabilizing effect of the secondary structure seems to be stronger than the stimulation of mRNA decay by polyadenylation. These findings and the detection of poly(A) addition sites within the stem-loop suggest a model with repeated cycles of polyadenylation and subsequent removal of the adenosines and adjacent nucleotides resulting in a stepwise degradation of the stem-loop structure and the rest of the RNA.

These results are, however, in contrast to other findings, in which a poly(A)-specific exoribonuclease seems not to be impeded by such a secondary structure in the potato *atp9* mRNA (Gagliardi *et al*., 2001).

The presence of non-encoded nucleotides at the 3' end of mRNAs raises the question of how these nucleotides are added. As already outlined in chapter 1.2, poly(A) tails are generally synthesized by poly(A) polymerases (PAP). However, such an enzyme has to date not been reported for plant mitochondria. A recent publication demonstrated that this protein from *Escherichia coli* incorporates nucleotides of all identities into the 3' extensions *in vitro* (Yehudai-Resheff and Schuster, 2000). In plant mitochondria, the presence of cytidines among the added nucleotides, especially the addition of CCA, raises the possibility, that tRNA terminal nucleotidyltransferase is involved in the respective process (Kuhn *et al*., 2001, Fig.1 and 7; Williams *et al*., 2000). This enzyme is together with PAP part of the nucleotidyl superfamily and may have the potential to add short extensions or even longer oligo(A) tracts (Steege, 2000).

3.3 Head to tail connected mRNAs are present in the plant mitochondrial steady state RNA

The analysis of pea *atp9* and *Oenothera atp1* mitochondrial transcripts by RT-PCR analysis revealed the presence of head to tail connected RNA molecules in plant mitochondria, anticipated from previous 3' RACE experiments (Kuhn *et al*., 2001). Similar observations were made during the investigation of polyadenylated potato atp9 RNAs. Here linear RNA species were found, in which a 5' fragment is fused to the mature 3' end of the same molecule (Gagliardi *et al*., 2001). The possibility that

such amplificated fragments derive from RT-PCR artifacts seems unlikely, since most of the 5' and 3' ends detected in the joined molecules coincide with previously mapped termini (Schuster *et al*., 1986; Morikami and Nakamura, 1993; Kuhn *et al*., 2001). In addition transcripts with head to tail conformation are not detectable for every gene, since no such RNAs were found among pea *cox2* transcripts. Transcripts with head to tail conformation represent only a minor fraction since two rounds of PCR with nested primers are necessary to amplify cDNAs from these RNAs.

In about a third of the cDNA clones deriving from *Oenothera atp1* head to tail transcripts, 3' ends are located in the upstream inverted repeat. Such 3' ends were not found in clones originating from RNA that has been treated with T4 RNA ligase or nuclease S1 protection experiments and thus seem to be restricted to the head to tail connected molecule conformation. Moreover, none of these clones contains any non-encoded nucleotides, while short extensions of up to five nucleotides are found attached to the *atp1* transcripts with 3' ends downstream of the second stem-loop. The reason for this bias is unclear.

Short extensions of predominantly non-encoded adenosines are also present at the pea *atp9* 5' to 3' linkage, which correlate in length with the extensions previously identified in the above mentioned 3' RACE analysis. However, in the latter substantially more cytidines are found. As a large proportion of the head to tail linked mRNAs of both genes investigated do not contain any non-encoded nucleotides demonstrating that the formation of this molecule form is independent from the addition of such nucleotides.

The analysis of head to tail linked *atp1* transcripts revealed a previously unidentified 5' end about 75 nt downstream of the transcription initiation site. A 21 nt long conserved sequence motif encompassing this terminus is also found around the 5' ends of mature 18S rRNA in different plant species (Kuhn and Binder, 2002, manuscript, Fig.6). The 5' ends of the rRNAs are found exactly at the same position as those of the *atp1* transcripts. Since the mature 5' end of 18S rRNAs is generated by an endonucleolytic processing, a similar scenario is likely for the new identified *atp1* 5' end. The function and significance of this processing event remains unclear, since only about half of the steady state *atp1* transcripts contain the processed ends as indicated by the primer extension experiment Binder, 2002, manuscript, Fig.5).

Therefore I conclude that this terminus is generated by processing rather than being a primary 5' terminus resulting from transcription initiation.

One of the most interesting questions is, how these head to tail joined molecules are generated. Several arguments favour an RNA ligase activity.

First, except for one clone all *atp1* 5' ends correspond to the new processed terminus. A similar bias is observed among the 5' ends detected after *in vitro* ligation with T4 RNA ligase (Kuhn and Binder, manuscript, Fig 4C). Thus head to tail connection *in vitro* and *in vivo* has a clear bias towards the shorter end. This preference for the shorter ends is determined by the ligase reaction mechanism, which allows only the ligation of 5' ends with monophosphate groups, suggesting that 5' to 3' end *in vivo* joining in mitochondria is performed by an RNA ligase. (Kuhn and Binder, 2002).

Second, no *in vivo* head to tail connected *cox2* transcripts are detectable in pea mitochondria. Here three different 5' termini are mapped and there is strong evidence that all of them originate from transcription initiation sites, thus carrying 5' triphosphates (Kuhn and Binder, 2002). Consistent with these 5' ends in cDNAs from *in vitro* ligated RNAs never coincide with the primary 5' termini mapped by primer extension analysis (Kuhn and Binder, 2002; cp. 2.4).Only a subpopulation of RNAs with slightly shortened 5' ends are ligated, resulting in an extremely low amount of *in vivo* head to tail connected RNA species escaping detection by RT-PCR.

Third, RNA ligases have recently been identified in mitochondria of the kinetoplastide *Trypanosoma brucei*, where they play a crucial role in post-transcriptional RNA editing processes (Huang *et al*., 2001). Here the RNA ligases are involved in the ligation of the mRNA moieties after insertion or deletion of uridine(s), but there exists no report of head to tail ligated RNAs in kinetoplastide. Nothing is known about RNA ligases in plant mitochondria, but my results on the linkage of mRNA termini point to a process requiring an RNA ligase in plant mitochondria.

Whatever the function of such head to tail connected mRNAs may be, they indicate a very close interaction of the 5' and 3' extremities, even if the formation of the phosphodiester bond occurs incidentally. Such close interactions have been shown for nuclear/cytoplasmic RNAs resulting in a circularized transcript (cp. 1.3). In this context circularization is mediated by RNA-protein and protein-protein interaction (Wells *et al*., 1998).

3.4 CR-RT-PCR is an appropriate method for the simultaneous identification of 5' and 3' transcript ends and the detection of non-encoded nucleotides

The CR-RT-PCR method was developed for the simultaneous investigation of both transcript extremities as well as of non-encoded nucleotides. The validity of this method should be tested by the analysis of pea *cox2* mRNAs that have been mapped previously (Moon *et al*., 1985). The results of the CR-RT-PCR were confirmed by primer extension, *in vitro* transcription and nuclease S1 mapping analyses (Kuhn and Binder, 2002). The primer extension analysis identified three new 5' ends each differing by several nucleotides from the already known ends (Kuhn and Binder, 2002, Fig.3). The discrepancies probably derive from the different techniques used. The sequences surrounding the two downstream 5' ends show significant similarities with a conserved nonanucleotide motif (CNM) characteristic for plant mitochondrial promoters (Binder *et al*., 1995; Dombrowski *et al*., 1999; Kuhn and Binder, 2002, Fig.7). Run-off products consistent with a transcription initiation at these sites confirmed the function of these sequences as promoters (Kuhn and Binder, 2002, Fig.4). In contrast, a completely different sequence surrounds the most upstream 5' initiation site. Nevertheless, weak signals corresponding to run-off products initiated at this site indicate a promoter activity of this sequence *in vitro*.

The 5' ends detected in the cDNAs derived from the CR-RT-PCR amplification revealed different 5' termini scattering over 32 nt downstream of the transcription start sites. This distribution is in contrast to the primer extension experiments with three major 5' ends prevailing and is most likely due to di- or triphosphate groups present at primary transcripts whose 5' end derive from *de novo* initiation events, which are most likely not a substrate for the T4 RNA ligase (Kuhn and Binder, 2002). The *cox2* mRNAs slightly shortened at the 5' end are possibly products of processing or degradation events by endo- or exonucleolytic activities rather than resulting from initiation events downstream of the transcription initiation sites. Thus these 5' ends are supposed to carry monophosphate groups suitable for ligation through RNA ligase. The hypothesis is further substantiated by the CR-RT-PCR analysis of *Oenothera atp1* transcripts, which identified almost exclusively 5' ends downstream of the previously mapped transcription initiation site (cp. 2.3). Thus, on the one hand the CR-RT-PCR analysis can be applied to detect 5' processed ends, on the other hand 5' termini deriving from *de novo* transcription initiation events can be detected

after a pre-treatment of the RNAs with consecutive phosphatase and kinase reactions (unpublished results). A similar experimental strategy has been already applied for the 5' analysis of capped nuclear transcripts after decapping of the respective RNAs by tobacco acid pyrophosphatase (Mandl *et al*., 1991). Moreover, the method has been successfully used for the analysis of land snail mitochondrial tRNAs (Yokobori and Paabo, 1995).

The CR-RT-PCR analysis identified a new 3' end of the *cox2* transcript located about 100 nt further downstream of the previously mapped end (Moon *et al*., 1986; Kuhn *et al*., 2002). In addition this new end was confirmed by a nuclease S1 analysis. Whether the *cox2* 3' end is formed directly by transcription termination or indirectly by post-transcriptional processes is unclear at present. Two alternative stem-loop structures can be folded a few nucleotides upstream of the slightly scattering 3' termini (Kuhn and Binder, 2002, Fig.8). These structures and the location of the 3' end within a polythymidine stretch of 9 nt resemble prokaryotic transcription terminators (Lesnik *et. al*., 2001). This suggests the *cox2* 3' termini originate from transcription termination, which of course has to be investigated in detail by *in vitro* studies.

The CR-RT-PCR analysis of the *cox2* mRNAs revealed no non-encoded nucleotides at the *cox2* 3' transcript termini (cp. 2.3). In addition a 3' RACE analysis with oligo(dT) adapter primer also failed to detect 3' non-encoded nucleotides (Dospil, Kuhn and Binder, unpublished results). This suggests that either no non-encoded nucleotides are added to the *cox2* 3' end or such extensions are rapidly removed.

With the help of additional data, e. g. the rough size of the transcripts, this method can be applied to determine the transcriptional and processed 5' and 3' ends of even those transcripts with complex transcription patterns. In addition, the presence of non-encoded nucleotides can be detected in the same reaction.

4 SUMMARY

In this thesis several aspects of mRNA processing in plant mitochondria were investigated. Mainly, modes and parameters of 3' RNA processing were analyzed both *in vitro* and *in vivo*. Furthermore a protein was characterized that may play a role in 3' mRNA processing in plant mitochondria.

This latter, an RNA helicase from *Arabidopsis thaliana,* exhibits high similarity to the yeast SUV3 protein, which is thought to be a component of a processing and degradation complex in yeast mitochondria. Several independent experimental approaches showed that this nuclear-encoded RNA helicase (AtSUV3) is located in *Arabidopsis thaliana* mitochondria. The AtSUV3 mRNA of about 2500 nt is assembled from 16 exons of a weakly expressed unique gene and the predicted protein has a calculated molecular weight of 63.6 kDa. The N-terminal domain of the protein contains motifs characteristic for DExH box RNA helicases and exhibits a low endogenous ATPase activity, which could be stimulated by mitochondrial RNA.

To determine the influence of posttranscriptional modifications on 3' end processing and mRNA stability in plant mitochondria, pea *atp9* and *Oenothera atp1* transcripts were investigated for the presence of 3' non-encoded nucleotides. Poly(A) tails were predominantly identified at previously mapped 3' ends. Functional studies in a pea mitochondrial *in vitro* processing system revealed rapid removal of the poly(A) extensions up to a stabilizing stem-loop structure, while the complete degradation of the RNA is only moderately stimulated. In contrast, polyadenylation promoted total degradation in transcripts without a stabilizing stem-loop structure. An alternative 3' RACE approach identified non-encoded extensions with a maximal length of 3 nucleotides consisting most frequently of adenosines and cytidines. Processing assays of substrates containing homopolymer stretches of various length showed that 3 adenosines have no impact on RNA degradation, while 10 or more adenosines substantially accelerate RNA decay. Thus polyadenylation can generally destabilize mRNAs, but the degradation promoting effect is almost annihilated by stabilizing stem-loop structures. Together these antagonistic actions result in the efficient formation of 3' processed and stable transcripts.

The characterization of the pea *atp9* and *Oenothera atp1* transcript termini identified 5' to 3' end connected molecules. The formation of such RNAs is independent from the addition of non-encoded nucleotides but requires processed 5' ends with 5'

monophosphate groups. The biological function of these head to tail connected RNA molecules in plant mitochondria is unclear.

To gain a more general insight into mRNA processing in plant mitochondria, the 5' and 3' extremities of the previously characterized pea *cox2* gene were analyzed exemplarily by a new experimental approach based on the *in vitro* circularization of the respective RNA molecules by T4 RNA ligase followed by an RT-PCR analysis. The validity of this method, which allows the simultaneous mapping of 5' and 3' termini as well as the detection of non-encoded nucleotides, was confirmed by primer extension analysis and nuclease S1 mapping.

5 REFERENCES

Alexander, C., N. Faber and P. Klaff (1998) Characterization of protein binding to the spinach chloroplast psbA mRNA 5' region. *Nucleic Acids Res.* 26: 2265-2272.

Allmang, C. and D. Tollervey (1998) The role of the 3' external transcribed spacer in yeast pre-rRNA processing. *J Mol Biol.* 278: 67-78.

Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W Miller and D. J. Lipman (1997) Gapped-BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389-3402.

Anderson S, A. T. Bankier, B. G. Barrell, M. H. de Bruijn, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P.H. Schreier, A. J. Smith, R. Staden and I. G. Young (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290: 457-465.

Anderson, J. S. J. and R. Parker (1996) RNA turnover: The helicase story unwinds. *Curr. Biol.* 6: 780-782.

Andersson, S. G. E., A. Zomrodibour, J. O. Andersson, T. Sicheritz-Ponten, A. C. M. Aslmark, R. M. Podowski, A. K. Näslund, A. S. Eriksson, H. H. Winkler and C. G. Kurland (1998) The genome sequence of *Rickettsia provazekii* and the origin of mitochondria. *Nature* 396: 133-140.

Aubert, D., C. Bisanz-Seyer and M. Herzog (1992) Mitochondrial *rps14* is a transcribed and edited pseudogene in *Arabidopsis thaliana*. *Plant Mol Biol.* 20: 1169-1174.

August, J.T., P.J. Ortiz and J. Hurwitz (1962) Ribonucleic acid-dependent ribonucleotide incorporation: I. Purification and properties of the enzyme. *J. Biol. Chem.* 237: 3786-3793.

Baginsky S, A. Shteiman-Kotler, V. Liveanu, S. Yehudai-Resheff, M. Bellaoui, R. E. Settlage, J. Shabanowitz, D. F. Hunt, G. Schuster and W. Gruissem (2001) Chloroplast PNPase exists as a homo-multimer enzyme complex that is distinct from the *Escherichia coli* degradosome. RNA 7: 1464-1475.

Barkan, A. (1988) Proteins encoded by a complex chloroplast transcription unit are each translated from both monocistronic and polycistronic mRNAs. *EMBO J.* 7: 2637-2644.

Beelman, C. A. and R. Parker (1995) Degradation of mRNA in eukaryotes. *Cell* 81: 179-183.

Bessarab, D. A., V. R. Kaberdin, C. Wei, G. Liou and S. Lin-Chao (1998) RNA components of *Escherichia coli* degradosome: Evidence for rRNA decay. *Proc. Natl. Acad. Sci. USA* 95: 3157-3161.

Bibb M.J., R. A. Van Etten, C. T. Wright, M. W. Walberg and D. A. Clayton (1981) Sequence and gene organization of mouse mitochondrial DNA. *Cell* 26: 167-180.

Bonen, L. and S. Bird (1988) Sequence analysis of the wheat mitochondrial *atp6* gene reveals a fused upstream reading frame markedly divergent N termini along plant ATP6 proteins. *Gene* 73: 47-56.

Binder, S., F. Hatzack and A. Brennicke (1995) A novel pea mitochondrial *in vitro* transcription system recognizes homologous and heterologous mRNA and tRNA promoters. *J. Biol. Chem*. 270: 22182-22189.

Birnstiel, M. L., M. Busslinger and K. Strub (1985) Transcription termination and 3' processing: The end is in sight. *Cell* 41: 349-359.

Blum, E., A. J. Carpousis and C. F. Higgins (1999) Polyadenylation promotes degradation of 3'-structured RNA by the *Escherichia coli* mRNA degradosome *in vitro*. *J. Biol. Chem*. 274: 4009-4016.

Brandt, P., M. Unseld, U. Eckert-Ossenkopp and A. Brennicke (1993) An *rps14* pseudogene is transcribed and edited in *Arabidopsis* mitochondria. *Curr. Genet.* 24: 330-336.

Brennicke, A., A. Marchfelder and S. Binder (1999) RNA editing. *FEMS Microbiol. Rev.* 23: 297-316.

Bruick, R. K. and S. P. Mayfield (1998) Processing of the psbA 5' untranslated region in *Chlamydomonas reinhardtii* depends upon factors mediating ribosome association. *J. Cell. Biol*. 143: 1145-1153.

Cao, G. J. and N. Sarkar (1992) Identification of the gene for an *Escherichia coli* poly(A) polymerase. *Proc. Natl. Acad. Sci. USA* 89: 10380-10384.

Cao G., J. Pogliano and N. Sarkar (1996) Identification of the coding region for a second poly(A) polymerase in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 93: 11580-11585.

Carpousis, A. J., G. Van Houwe, C. Ehretsmann and H. M. Krisch (1994) Copurification of *E. coli* RNase E and PNPase: Evidence for a specific association between two enzymes important in RNA processing and degradation. *Cell* 76: 889-900.

Carpousis, A. J., N. F. Vanzo and L. C. Raynal (1999) mRNA degradation. A tale of poly(A) and multiprotein machines. *Trends Genet.* 15: 24-28.

Chen, H. C. and D. B. Stern (1991a) Specific binding of chloroplast proteins *in vitro* to the 3' untranslated region of spinach chloroplast *petD* mRNA. *Mol. Cell. Biol*. 11: 4380-4388.

Chen, H. C. and D. B. Stern (1991b) Specific ribonuclease activities in spinach chloroplasts promote mRNA maturation and degradation. *J. Biol. Chem.* 266: 24205-24211.

Chen, Q., C. C. Adams, L. Usack, J. Yang, R. A. Monde and D. B. Stern (1995) An AU-rich element in the 3' untranslated region of the spinach chloroplast *petD* gene participates in sequence-specific RNA-protein complex formation. *Mol. Cell. Biol*. 15: 2010-2018.

Coburn, G. A. and G. A. Mackie (1998) Reconstitution of the degradation of the mRNA for ribosomal protein S20 with purified enzymes. *J. Mol. Biol*. 279: 1061-1074.

Cohen, S. N. (1995) Surprises at the 3' end of prokaryotic RNA. *Cell* 80: 829-832.

Conrad-Webb, H., P. S. Perlman, Z. Hong and R. A. Butow (1990) The nuclear SUV3-1 mutation affects a variety of post-transcriptional processes in yeast mitochondria. *Nucleic Acids Res.* 18: 1369-1376.

de Duve, C. (1996) Die Herkunft der komplexen Zellen. *Spektrum der Wissenschaft*, Juni, 60-68.

Dell'Orto, P., A. Moenne, P. V. Graves and X. Jordana (1993) The potato mitochondrial ATP synthase subunit 9: gene structure, RNA editing and partial protein sequence. *Plant Science* 88: 45-53.

Deutscher, M. P. and N. B. Reuven (1991) Enzymatic basis for hydrolytic versus phosphorolytic mRNA degradation in *Escherichia coli* and *Bacillus subtilis. Proc. Natl. Acad. Sci. USA* 88: 3277-3280.

Dewey, R. E., C. S. 3rd Levings and D. H. Timothy (1986) Novel recombinations in maize mitochondrial genome produce a unique transcriptional unit in the texas male sterile cytoplasm. *Cell* 44: 439-449.

Dombrowski, S., A. Brennicke and S. Binder (1997) 3'-inverted repeats in plant mitochondrial mRNAs are processing signals rather than transcription terminators. *EMBO J.* 16: 5069-5076.

Dombrowski, S., M. Hoffmann, J. Kuhn, A. Brennicke and S. Binder (1998) On the mitochondrial promoters in *Arabidopsis thaliana* and other flowering plants. In: Plant Mitochondria: From Gene to Function, 165-170.

Dombrowski, S., M. Hoffmann, C. Guha and S. Binder (1999) Continuous primary sequence requirements in the 18-nucleotide promoter of dicot plant mitochondria. *J. Biol. Chem.* 274: 10094-10099.

Doolittle, W. F. (1998) A paradigm gets shifty. *Nature* 392: 15-16.

Drager, R. G., M. Zeidler, C. L. Simpson and D. B. Stern (1996) A chloroplast transcript lacking the 3' inverted repeat is degraded by 3' to 5' exoribonuclease activity. *RNA* 2: 652-663.

Drager, R. G., J. Girard-Bascou, Y. Choquet, K. L. Kindle and D. B. Stern (1998) *In vivo* evidence for 5'->3' exoribonuclease degradation of an unstable chloroplast mRNA. *Plant J.* 13: 85-96.

Drager, R. G., D. C. Higgs, K. L. Kindle and D. B. Stern (1999) 5' to 3' exoribonucleolytic activity is a normal component of chloroplast mRNA decay pathways. *Plant J.* 19: 521-531.

Feng, Y., H. Huang, J. Liao and S. N. Cohen (2001) *Escherichia coli* poly(A) -binding proteins that interact with components of degradosomes or impede RNA decay mediated by polynucleotide phosphorylase and RNase E. *J. Biol. Chem.* 276: 31651-31656.

Flaherty S. M., P. Fortes, E. Izaurralde, I. W. Mattaj and G. M. Gilmartin (1997) Participation of the nuclear cap binding complex in pre-mRNA 3' processing. *Proc Natl Acad Sci USA* 94: 11893-11898.

Gagliardi, D. and C. J. Leaver (1999) Polyadenylation accelerates the degradation of the mitochondrial mRNA associated with cytoplasmic male sterility in sunflower. *EMBO J.* 18: 3757-3766.

Gagliardi, D., J. Kuhn, U. Spadinger, A. Brennicke, C. J. Leaver and S. Binder (1999) An RNA helicase (AtSUV3) is present in *Arabidopsis thaliana* mitochondria. *FEBS Lett.* 458: 337-442.

Gagliardi, D., R. Perrin, L. Maréchal-Drouard, J.-M. Grienenberger and C. J. Leaver (2001) Plant mitochondrial polyadenylated mRNAs are degraded by a 3'-to 5'-exoribonuclease activity, which proceeds unimpeded by stable secondary structures *J. Biol. Chem*. 276: 43541-43547.

Ghosh, S. and M. P. Deutscher (1999) Oligoribonuclease is an essential component of the mRNA decay pathway. *Proc. Natl. Acad. Sci. USA* 96: 4372-4377.

Giegé, P. and A. Brennicke (1999) RNA editing in *Arabidopsis* mitochondria effects 441 C to U changes in ORFs. *Proc. Natl. Acad. Sci. USA* 96: 15324-15329.

Giegé, P., M. Hoffmann, S. Binder and A. Brennicke (2000) RNA degradation buffers asymmetries of transcription in *Arabidopsis* mitochondria. *EMBO Rep.* 1: 164-170.

Gray, M. W. (1989) The evolutionary origins of organelles. *Trends Genet*. 5: 294-299.

Gray, M. W. (1992) The endosymbiont hypothesis revisited. *Int. Rev. Cytol*. 141: 233-357.

Gray, M. W. (1998) *Rickettsia*, typhus and mitochondrial connection. *Nature* 396: 109-110.

Gray, M.W. (1999a) Evolution of organellar genomes. *Curr. Opin. Genet. Dev*. 9: 678-687.

Gray, M. W. (1999b) Mitochondrial evolution. *Science* 283: 1476-1481.

Gualberto, J. M., C. Demon, J.-H. Weil and J.-M. Grienenberger (1990) Structure and transcription of the gene coding for subunit 3 of cytochrome oxidase in wheat mitochondria. *Curr. Genet.* 17: 41-47.

Haff, L. A. and L. Bogorad (1976) Poly(adenylic-acid)-containing RNA from plastids of maize. *Biochemistry* 15: 4110-4115.

Hajnsdorf, E., F. Braun, J. Haugel-Nielsen and P. Régnier (1995) Polyadenylation destabilizes the *rpsO* mRNA of *Escherichia coli. Proc. Natl. Acad. Sci. USA* 92: 3973-3977.

Hayes, R., J. Kudla, G. Schuster, L. Gabay, P. Maliga and W. Gruissem (1996) Chloroplast mRNA 3'-end processing by a high molecular weight complex is regulated by nuclear encoded RNA binding proteins. *EMBO J.* 15: 1132-1142.

Hedtke, B., A. Weihe and T. Börner (1997) Mitochondrial and chloroplast phage-type RNA polymerases in *Arabidopsis. Science* 277: 809-811.

Hoffmann, M., S. Dombrowski, C. Guha and S. Binder (1999) Cotranscription of the *rpl5-rps14-cob* gene cluster in pea mitochondria. *Mol. Gen. Genet*. 261: 537-545.

Huang, C. E., J. Cruz-Reyes, A. G. Zhelonkina, S. O´Hearn, E. Wirtz and B. Sollner-Webb (2001) Roles for ligases in the RNA editing complex of *Trypanosoma brucei*: bandIV is needed for U-deletion and RNA repair. *EMBO J*. 20: 4694-4704.

Jäger, S., O. Fuhrmann, C. Heck, M. Hebermehl, E. Schiltz, R. Rauhut and G. Klug (2001) An mRNA degrading complex in *Rhodobacter capsulatus*. *Nucleic Acidc Res*. 29: 4581-4588.

Kaleikau, E. K., C. P. Andre and V. Walbot (1992) Structure and expression of rice mitochondrial apocytochrome b gene (*cob-1*) and pseudo-gene (*cob-2*). *Curr. Genet.* 22: 463-470.

Kim, M., D. A. Christopher and J. E. Mullet (1993) Direct evidence for selective modulation of psbA, rpoA, rbcL and 16S rRNA stability during barley chloroplast development. *Plant Mol. Biol.* 22: 447-463.

Klaff, P. (1995) mRNA decay in spinach chloroplasts: psbA mRNA degradation is initiated by endonucleolytic cleavages within the coding region. *Nucleic Acids Res*. 23: 4885-4892.

Kubo, T., S. Nishizawa, A. Sugawara, N. Itchoda, A. Estiati and T. Mikami (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.

Kudla, J., R. Hayes and W. Gruissem (1996) Polyadenylation accelerates degradation of chloroplast mRNA. *EMBO J.* 15: 7137-7146.

Kuhn, J., U. Tengler and S. Binder (2001) Transcript lifetime is balanced between stabilizing stem-loop structures and degradation promoting polyadenylation in plant mitochondria. *Mol. Cell. Biol.* 21: 731-742.

Kuhn, J. and S. Binder (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.

Kunzmann, A., A. Brennicke and A. Marchfelder (1998) 5' end maturation and RNA editing have to precede tRNA 3' end processing in plant mitochondria. *Proc. Natl. Acad. Sci. USA* 95: 108-113.

Lang, B. F., G. Burger, C. J. O´Kelly, R. Cedergren, G. B. Golding, C. Lemieux, D. Sankoff, M. Turmel and M. W. Gray (1997) An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* 387: 493-497.

Lee, K and S. N. Cohen (2001) Effects of 3' terminus modifications on mRNA functional decay during *in vitro* protein synthesis. *J. Biol. Chem.* 276: 23268-23274.

Lesnik, E. A., R. Sampath, H. B. Levene, T. J. Henderson, J. A. McNeil and D. J. Ecker (2001) Prediction of rho-independent transcriptional terminators in *Escherichia coli. Nucleic Acids Res*. 29: 3583-3594.

Li, Z., S. Pandit and M. P. Deutscher (1998) Polyadenylation of stable RNA precursors *in vivo*. *Proc. Natl. Acad. Sci. USA* 95: 12158-12162.

Li, Z., S. Pandit and M. P. Deutscher (1999) RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. *EMBO J*. 18: 2878-2885.

Liou, G. G., W. N. Jane, S. N. Cohen, N. S. Lin and S. Lin-Chao (2001) RNA degradosomes exist *in vivo* in *Escherichia coli* as multicomponent complexes associated with the cytoplasmic membrane via the N-terminal region of ribonuclease E. *Proc. Natl. Acad. Sci. USA* 98: 63-68.

Lisitsky I., P. Klaff and G. Schuster (1996) Addition of destabilizing poly (A)-rich sequences to endonuclease cleavage sites during the degradation of chloroplast mRNA. *Proc. Natl. Acad. Sci. USA* 93: 13398-13403.

Lisitsky, I., P. Klaff and G. Schuster (1997a) Blocking polyadenylation of mRNA in the chloroplast inhibits its degradation. *Plant Journal* 12: 1173-1178.

Lisitsky, I., A., Kotler and G. Schuster (1997b) The mechanism of preferential degradation of polyadenylated RNA in the chloroplast. The exoribonuclease 100RNP/polynucleotide phosphorylase displays high binding affinity for poly(A) sequences. *J. Biol. Chem.* 272: 17648-17653.

Liu, A. W., K. K. Narayanan, C. P. André, E. K. Kaleikau and V. Walbot (1992) Cotranscription of *orf25* and *cox3* in rice mitochondria. *Curr. Genet*. 21: 507-513.

Lupold, D. S., A. G. F. S. Caoile, and D. B. Stern (1999) Polyadenylation occurs at multiple sites in maize mitochondrial *cox2* mRNA and is independent of editing status. *Plant Cell* 11: 1565-1577.

Mackie, G. A. (1998) Ribonuclease E is a 5'-end-dependent endonuclease. *Nature* 395: 720-723.

Mackie, G. A. (2000) Stabilization of circular rpsT mRNA demonstrates the 5'-end dependence of RNase E action *in vivo*. *J. Biol. Chem*. 275: 25069-25072.

Mandl, C. W., F. X. Heinz, E. Puchhammer-Stockl and C. Kunz (1991) Sequencing the termini of capped viral RNA by 5'-3' ligation and PCR. *Biotechniques*, 10: 484-486.

Margossian, S. P. and R. A. Butow (1996) RNA turnover and the control of mitochondrial gene expression. *Trends Biochem. Sci.* 21: 392-396.

Margossian, S. P., H. Li, H. P. Zassenhaus and R. A. Butow (1996) The DExH box protein Suv3p is a component of a yeast mitochondrial 3'-to-5' exoribonuclease that suppresses group I intron toxicity. *Cell* 84: 199-209.

Margulis L. (1970) Origin of eukaryotic cells. Yale University Press, New Haven, CT.

Martin W. and M. Müller (1998) The hydrogen hypothesis for the first eukaryote. *Nature* 392: 37-41.

McLaren, J. L., S. F. Newbury, G: S. C. Dance, H. C. Causton and C. F: Higgins (1991) mRNA degradation by processive 3'-5' exoribonucleases *in vitro* and the implications for procaryotic mRNA decay *in vivo. J. Mol. Biol.* 221: 81-95.

Michel, F., K. Umesono and H. Ozeki (1989) Comparative and functional anatomy of group II catalytic introns - a review. *Gene* 82: 5-30.

Miczak, A., V. R. Kaberdin, C. L. Wie and S. Lin-Chao (1996) Proteins associated with RNase E in a multicomponent ribonucleolytic complex. *Proc. Natl. Acad. Sci. USA* 93: 3865-3869.

Min, J. and H. P. Zassenhaus (1993) Identification of a protein complex that binds to a dodecamer sequence found at the 3' ends of yeast mitochondrial mRNAs. *Mol. Cell. Biol.* 13: 4167-4173.

Mitchell, P and D. Tollervey (2000a) mRNA stability in eukaryotes. *Curr. Opin. Genet. Dev.* 10: 193-198.

Mitchell, P. and D. Tollervey (2000b) Musing on the structural organization of the exosome complex. *Nature Struct. Biol.* 7: 843-846.

Mitchell, P. and D. Tollervey (2001) mRNA turnover. *Curr. Opin. Cell Biol.* 13: 320-325.

Mitchell, P., E. Petfalski, A. Shevchenko, M. Mann und D. Tollervey (1997). The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'->5' exoribonucleases. *Cell* 91: 457-466.

Mohanty B. K. and S. R. Kushner (1999) Analysis of the function of *Escherichia coli* poly(A) polymerase I in RNA metabolism. *Mol. Microbiol.* 34: 1094-1108.

Mohanty, B. K. and S. R. Kushner (2000a) Polynucleotide phosphorylase, RNase II and RNase E play different roles in the *in vivo* modulation of polyadenylation in *Escherichia coli*. *Mol. Microbiol*. 36: 982-994.

Mohanty, B. K. and S. R. Kushner (2000b) Polynucleotide phosphorylase functions both as a 3'->5' exonuclease and a poly(A) polymerase in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 97: 11966-11971.

Monde, R. A., G. Schuster and D. B. Stern (2000) Processing and degradation of chloroplast mRNA. *Biochimie* 82: 573-582.

Moon, E., T. H. Kao and R. Wu (1985) Pea cytochrome oxidase subunit II gene has no intron and generates two mRNA transcripts with different 5' termini. *Nucleic Acids Res*. 13: 3185-3212.

Moreira, D., H. Le Guyader and H. Phillippe (2000) The origin of red algae and the evolution of chloroplasts. *Nature* 405: 32-33.

Morikami, A. and K. Nakamura (1987) Structure and expression of pea mitochondrial F1ATPase alpha-subunit gene and its pseudogene involved in homologous recombination. *J. Biochem.* 101: 967-976.

Morikami, A. and K. Nakamura (1993) Transcript map of oppositely oriented pea mitochondrial genes encoding the α-subunit and the subunit 9 of F_0F_1-ATPase complex. *Biosci. Biotechnol. Biochem*. 57: 1530-1535.

Müller, M. 1993 The hydrogenosome. *J Gen. Microbiol*. 139: 2879-89.

Nickelsen, J., M. Fleischmann, E. Boudreau,. M. Rahire and J. D. Rochaix (1999) Identification of cis-acting RNA leader elements required for chloroplast *psbD* gene expression *Chlamydomonas*. *Plant Cell* 11: 957-970.

Nudler, E., M. Kashlev, V. Nikiforov and A. Goldfarb (1995) Coupling between transcription termination and RNA polymerase inchworming. *Cell* 81: 351-357.

O'Hara, E. B., J. A. Chekanova, C. A. Ingle, Z. R. Kushner, E. Peters and S. R. Kushner (1995) Polyadenylation helps regulate mRNA decay in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 92: 1807-1811.

Oda, K., E. Yamato, E. Ohta, Y. Nakamura, M. Takemura, N. Nozato, K. Akashi, T. Kanegae, Y. Ogura, T. Kohchi and K. Ohyama (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. *J. Mol. Biol.* 223: 1-7.

Palmer, J. D. (2000) A single birth of all plastids? *Nature* 405: 69-72.

Palmer, J. D. and L. A. Herborn (1987) Unicircular structure of the *Brassica hirta* mitochondrial genome. *Curr. Genet*. 11: 565-570.

Proudfoot, N. J. (2001) Genetic dangers in poly(A) signals. *EMBO Rep*. 2: 891-892.

Py, B., C. F. Higgings, H. M. Krisch and A. J. Carpousis (1996) A DEAD-box RNA helicase in the *Escherichia coli* RNA degradosome. *Nature* 381: 169-172.

Quiñones, V., S. Zanlungo, A. Moenne, I. Gomez, L. Holuige, S. Litvak and X. Jordana (1996) The *rpl5-rps14-cob* gene arrangement in *Solanum tuberosum*: rps14 is a transcribed and unedited pseudogene. *Plant Mol. Biol.* 31: 937-943.

Rapp, W. D. and D. B. Stern (1993) A conserved 11 nucleotide sequence contains an essential promoter element of the maize mitochondrial *atp1* gene. *EMBO J*. 11: 1065-1073.

Rapp, W. D., D. S. Lupold, S. Mack and D. B. Stern (1993) Architecture of the maize mitochondrial atp1 promoter as determined by linker scanning and point mutagenesis. *Mol. Cell. Biol.* 13: 7232-7238.

Reinbothe, S., C. Reinbothe, C. Heintzen, C. Seidenbecher and B. Parthier (1993) A methyl jasmonate-induced shift in the length of the 5' untranslated region impairs translation of the plastid *rbcL* transcript in barley. *EMBO J*. 12: 1505-1512.

Rott, R., R. G. Drager, D. B. Stern and G. Schuster (1996) The 3' untranslated regions of chloroplast genes in *Chlamydomonas reinhardtii* do not serve as efficient transcriptional terminators. *Mol. Gen. Genet*. 252: 676-683.

Rott, R., V. Liveanu, R. G. Drager, D. B. Stern and G. Schuster (1998a) The sequence and structure of the 3' untranslated regions of chloroplast transcripts are important determinants of mRNA accumulation and stability. *Plant mol. biol*. 36: 307-314.

Rott, R., H. Levy, R. G. Drager, D. B. Stern and G. Schuster (1998b) 3'-processed mRNA is preferentially translated in *Chlamydomonas reinhardtii. Mol. Cell. Biol*. 18: 4605-4611.

Sakamoto, W., N. R. Sturm, K. L. Kindle and D. B. Stern (1994) *petD* mRNA maturation in *Chlamydomonas reinhardtii* chloroplasts: role of 5' endonucleolytic processing. *Mol. Cell. Biol*. 14: 6180-6186.

Saalaoui, E., S. Litvak and A. Araya (1990) The apocytochrome b from an alloplasmic line of wheat (*T. aestivum, cytoplasm-T. timopheevi*) exists in two differently expressed forms. *Plant Sci*. 66: 237-246.

Shapira, M., A. Lers, P. B. Heifetz, V. Irihimovitz, C. B. Osmond, N. W. Gillham und J. E. Boynton (1997) Differential regulation of chloroplast gene expression in *Chlamydomonas reinhardtii* during photoacclimation: light stress transiently suppresses synthesis of the Rubisco LSU protein while enhancing synthesis of the PS II D1 protein. *Plant Mol Biol* 33: 1001-1011.

Schimper, A. F. W. (1883) Über die Entwicklung der Chlorophyllkörner und Farbkörner. *Bot. Zeit.* 41:105-114.

Schuster, W., R. Hiesel, P. G. Isaac, C. J. Leaver and A. Brennicke (1986) Transcript termini in messenger RNAs in higher plant mitochondria. *Nucleic Acids Res*. 14: 5943-5954.

Schuster, G. and W. Gruissem (1991) Chloroplast mRNA 3' end processing requires a nuclear-encoded RNA-binding protein. *EMBO J*. 10: 1493-1502.

Steege, D. A. (2000) Emerging features of mRNA decay in bacteria. *RNA* 6: 1079-1090.

Stepien, P. P., S. P. Margossian, D. Landsman and R. Butow (1992) The yeast nuclear gene *suv3* affecting mitochondrial post-transcriptional processes encodes a putative ATP-dependent RNA-helicase. *Proc. Natl. Acad. Sci. USA* 89: 6813-6817.

Stern, D. B. and W. Gruissem (1987) Control of plastid gene expression: 3' inverted repeats act as mRNA processing and stabilizing elements, but do not terminate transcription. *Cell* 51: 1145-1157.

Stern, D. B., H. Jones and W. Gruissem (1989) Function of plastid mRNA 3' inverted repeats. RNA stabilization and gene-specific protein binding. *J. Biol. Chem*. 264: 18742-18750.

Stern, D. B. and K. L. Kindle (1993) 3' end maturation of the *Chlamydomonas reinhardtii* chloroplast *atpB* mRNA is a two-step process. *Mol. Cell. Biol*. 13: 2277-2285.

The *Arabidopsis* Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408: 796-815.

Tomine, Y., L. Kwon, M. C. Anguera, G. Schuster and D. B. Stern (2000) Polyadenylation of three classes of chloroplast RNAs in *Chlamydomonas reinhardtii. RNA* 6: 598-607.

Travis, J. (1998) The hydrogen hypothesis: How did complex cells get their power stations? *Science News* 153: 253-255.

Unseld, M., J. R. Marienfeld, P. Brandt and A. Brennicke (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nature Genet*. 15: 57-61.

Vaistij, F. E., M. Goldschmidt-Clermont, K. Wostrikoff and J. D. Rochaix (2000) Stability determinants in the chloroplast *psbB/T/H* mRNAs in *Chlamydomonas reinhardtii*. *Plant J*. 21: 469-482.

Vanzo, N. F., Y. S. Li, B. Py, E. Blum, C. F. Higgins, L. C. Raynal, H. M. Krisch and A. J. Carpousis (1998) Ribonuclease E organizes the protein interactions in the *Escherichia coli* RNA degradosome. *Genes Dev.* 12: 2770-2781.

Wahleithner, J. A. and D. R. Wolstenholme (1988) Ribosomal protein S14 genes in broad bean mitochondrial DNA. *Nucleic Acids Res*. 16: 6897.6913.

Wallin, I. E. (1927) Symbioticism and the origin of species. Williams and Wilkins, Baltimore

Wang, Z. and M. Kiledjian (2001) Functional Link between the mammalian exosome and mRNA decapping. *Cell* 107: 751-762.

Ward, B. L., R. S. Anderson and A. J. Bendich (1981) The mitochondrial genome is large and variable in a family of plants (*Cucurbitaceae*). *Cell* 25: 793-803.

Wells S.E., P. E. Hillner, R. D. Vale, A. B. Sachs (1998) Circularization of mRNA by eukaryotic translation initiation factors. *Mol. Cell*. 2: 135-140.

Westhoff, P. und R. G. Herrmann (1988) Complex RNA maturation in chloroplasts. The *psbB* operon from spinach. *Eur. J. Biochem.* 171: 551-564.

Williams, M. A., Y. Johzuka and R. M. Mulligan (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.

Xu, F. and S. N. Cohen (1995) RNA degradation in *Escherichia coli* is regulated by 3' adenylation and 5' phosphorylation. *Nature* 374: 180-183.

Yang, J. and D. B. Stern (1997) The spinach chloroplast endoribonuclease CSP41 cleaves the 3'-untranslated region of *petD* mRNA primarily within its terminal stem-loop structure. *J. Biol. Chem*. 272: 12874-12880.

Ye, F., J. Bernhardt and W. O. Abel (1993) Genes for ribosomal proteins S3, L16, L5 and S14 are clustered in the mitochondrial genome of *Brassica napus* L.. *Curr. Genet*. 24: 323-329.

Yehudai-Resheff, S. and G. Schuster (2000) Characterization of the *E. coli* poly(A) polymerase: nucleotide specificity, RNA binding affinities and RNA structure dependence. *Nucleic Acids Res*. 28: 1139-1144.

Yehudai-Resheff, S., M. Hirsh and G. Schuster (2001) Polynucleotide phosphorylase functions as both an exonuclease and a poly(A) polymerase in spinach chloroplasts. *Mol. Cell. Biol*. 21: 5408-5416.

Yokobori, S. and S. Paabo (1995) Transfer RNA editing in land snail mitochondria. *Proc. Natl. Acad. Sci. USA* 92: 10432-10435.

Yuckenberg, P. D. and S. L. Phillips (1982) Oligoadenylate is present in the mitochondrial RNA of *Saccharomyces cerevisiae*. *Mol. Cell. Biol*. 2: 450-456.

6 APPENDIX

6.1 Presentation of the results

6.1.1 An mRNA helicase (AtSUV3) is present in *Arabidopsis thaliana* mitochondria

FEBS 22610 FEBS Letters 458 (1999) 337–342

An RNA helicase (AtSUV3) is present in *Arabidopsis thaliana* mitochondria

Dominique Gagliardi[a], Josef Kuhn[b], Ulrike Spadinger[b], Axel Brennicke[b], Christopher J. Leaver[a], Stefan Binder[b,*]

[a] *Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK*
[b] *Allgemeine Botanik, Universität Ulm, D-89069 Ulm, Germany*

Received 3 August 1999

Abstract The proteins involved in mitochondrial mRNA processing and degradation in higher plants have yet to be identified. As a first step towards this aim, we report here the characterisation of a nuclear-encoded DExH box RNA helicase (AtSUV3) localised in *Arabidopsis thaliana* mitochondria. The AtSUV3 mRNA is assembled from the 16 exons of a weakly expressed unique gene and the predicted protein has a calculated molecular weight of 63.6 kDa. Subcellular fractionation of transgenic plants expressing AtSUV3/GUS fusion proteins localises this protein in mitochondria. The N-terminal domain of AtSUV3 containing the motifs characteristic of DExH box RNA helicases exhibits a low endogenous ATPase activity in vitro which can be stimulated by the presence of mitochondrial RNA, confirming that AtSUV3 is an RNA helicase.

© 1999 Federation of European Biochemical Societies.

Key words: Plant mitochondrion; RNA metabolism; RNA helicase; RNA processing

1. Introduction

In plant organelles, processing of mRNA molecules and the rate of RNA degradation are major contributors to regulating overall as well as individual gene expression. These processing steps include such diverse reactions as intron splicing, RNA editing, separation of individual coding regions from multicistronic RNA precursors and 5′ and 3′ trimming of mRNAs. Particularly the processing steps at the 3′ terminus of the RNA molecules are often only one step ahead of – and thus intimately connected to – ultimate degradation.

In both chloroplasts and plant mitochondria, potential hairpin structures present at the 3′ end of some transcripts are involved in the correct processing of 3′ termini [1,2]. In plastids such stem-loop structures interact with several proteins which generate mature 3′ termini by a combination of endonucleolytic and exonucleolytic processing reactions [3]. The secondary structures also interact with proteins that protect the RNA against ribonucleases and thus stabilise the mature RNA [3]. Exoribonucleolytic processing is probably also involved in the formation of the 3′ termini of transcripts in plant mitochondria [2]. Another feature shared between chloroplasts and plant mitochondria and also with bacteria is the surprising acceleration of RNA degradation by post-transcriptional polyadenylation [4–7].

In bacteria and chloroplasts, degradation of transcripts is often initiated by endoribonucleolytic cleavages and the resulting RNA fragments are then further degraded by exoribonuclease(s) [4,5,8,9]. Secondary structures present in the RNA that would otherwise impede the 3′ to 5′ progression of exonucleases can be resolved by RNA helicases. Indeed, in eubacteria such as *Escherichia coli* an RNA helicase (RhlB) is an essential part of the RNA degradosome which also contains polynucleotide phosphorylase (PNPase) and RNase E as major components [10]. Similarly in yeast mitochondria an RNA helicase (SUV3) is a component of a processing and degradation complex termed mtEXO with exoribonuclease activity [11–14]. The exoribonuclease activity is attributed to a putative RNase (DSS1) with similarity to the prokaryotic exoribonuclease RNase II [12].

In this report, we characterise an RNA helicase from *Arabidopsis thaliana* which exhibits high similarity to the yeast SUV3 protein and hence is termed AtSUV3. We show that the AtSUV3 protein is localised in mitochondria and exhibits functional properties of DExH RNA helicases. this identification of AtSUV3 is a first step towards further characterisation of the plant mitochondrial mRNA processing/degradation machinery.

2. Materials and methods

2.1. Identification of suv3 homologue in Arabidopsis thaliana (Atsuv3)

A suv3 homologous sequence was identified on chromosome IV of *A. thaliana* [15]. Specific PCR primers were used to amplify the predicted coding region from first strand cDNA of total RNA from *A. thaliana*. PCR products were cloned and seven individual clones were sequenced to obtain unambiguous sequence information. Sequencing was performed with a Thermo Sequenase fluorescent label kit (Amersham) and Cy5 AutoRead sequencing kits with Cy5-dATP labelling mix (Pharmacia). Sequencing products were detected and processed by an Alf Express sequencer (Pharmacia). The genomic reading frame predictions were extended and corrected by 3′-RACE experiments, which were performed as previously described [16]. Computer analyses were performed using Blast algorithms at the NCBI server. Primer extension analyses were performed with 3 μg poly(A)$^+$ RNA following previously described procedures [17]. Southern and Northern blot analyses were performed by standard procedures using a double-stranded probe representing the 5′ terminal two thirds of the Atsuv3 gene.

2.2. Subcellular localisation of the AtSUV3-GUS fusion protein

About two thirds of the Atsuv3 cDNA encoding amino acids 1–387 fused to the uidA (GUS) gene were cloned downstream of the 35S-CaMV promoter in the pBIN+ vector. The resulting plasmid pBIN+ Atsuv3-GUS was transformed into *Agrobacterium tumefaciens* CV3101. Wild-type Columbia ecotype *A. thaliana* plants and *Solanum tuberosum* (var. Desiree) plants were transformed as described by Bent and Clough [18] and by Sheerman and Bevan [19], respectively. Mitochondria were purified from transgenic *A. thaliana* and potato

*Corresponding author.

0014-5793/99/$20.00 © 1999 Federation of European Biochemical Societies. All rights reserved.
PII: S0014-5793(99)01168-0

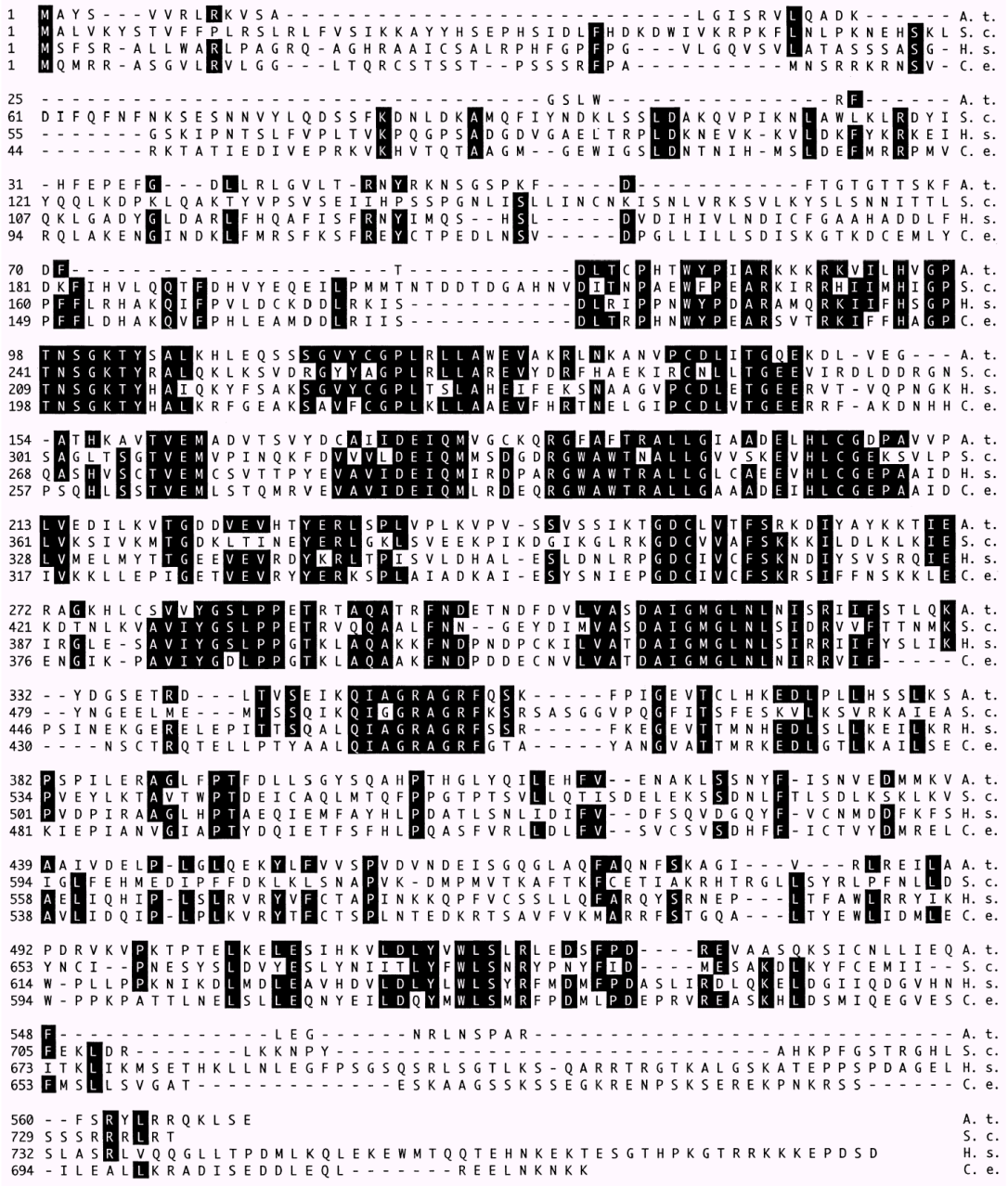


Fig. 1. The amino acid sequence deduced from the Atsuv3 cDNA (A. t.) is aligned with homologous protein sequences from yeast (S. c.), man (H. s.) and *Caenorhabditis* (C. r.). The conserved amino acids highlighted by inverse shading are conserved between at least three species and are concentrated in the central regions of the aligned homologues.

plants on sucrose density gradients as described by Leaver et al. [20]. The GUS activity of sucrose gradient fractions was determined by a fluorogenic assay as described by Jefferson [21]. Aliquots of sucrose gradient fractions were subjected to Western blot analysis by standard techniques using a monoclonal antibody directed against the mitochondrial E1α subunit of pyruvate dehydrogenase [22]. Immunoblots were developed with the Renaissance chemiluminescence reagent (NEN).

2.3. Overexpression and purification of His-tagged AtSUV3

A fragment of the Atsuv3 cDNA encoding amino acids 11–387 was cloned into the pQE32 plasmid (Qiagen) to generate the plasmid pQE32-Atsuv3. A culture of XL1-Blue cells transformed with the pQE32-Atsuv3 plasmid was grown, induced and harvested according to the manual. Bacteria were disrupted by sonication in 20 mM Tris-HCl pH 8.0, 100 mM NaCl (buffer TS) and inclusion bodies were pelleted by centrifugation of the bacterial lysate at 12 000×*g* for 10 min and resuspended in buffer TS containing 8 M urea. Insoluble material was removed by centrifugation at 20 000×*g* for 15 min and the supernatant was applied to TALONspin columns (Clontech). After two washes with buffer TS containing 10 mM imidazole, the His-tagged AtSUV3 protein was eluted with buffer TS containing 80 mM imidazole and dialysed for 3 h against 20 mM Tris-HCl pH 7.4, 50 mM NaCl, 3 mM $MgCl_2$, 1 mM DTT at 4°C. Most of His-tagged AtSUV3 precipitated during dialysis and was removed by centrifugation at 20 000×*g* for 10 min. Aliquots of the supernatant containing soluble His-tagged Atsuv3 were frozen at −20°C.

2.4. ATP hydrolysis assay

ATP hydrolysis activity was measured by incubating 222 GBq/

mmol [γ-^{32}P]ATP in the presence of soluble His-tagged AtSUV3 protein (100 ng) at 25°C for 4 h in a final volume of 20 µl of 20 mM Tris-HCl pH7.4, 50 mM NaCl, 3 mM $MgCl_2$, 1 mM DTT. Ten units of DNase I, 10 µg of RNase A, 500 ng of yeast transfer RNA or 500 ng of total mitochondrial RNA were added to the reaction. Aliquots of 1 µl were analysed by thin-layer chromatography on polyethyleneimine cellulose plates (Macherey-Nagel) in 1 M formic acid, 0.5 M LiCl and the reaction products were visualised by autoradiography.

3. Results

3.1. Structure of the Atsuv3 gene in A. thaliana

One of the genes predicted from the genomic sequence of a large part of chromosome IV of *A. thaliana* [15] has significant similarity with the yeast suv3 gene, which encodes a mitochondrial RNA helicase [23]. thisis potential plant protein shares 39% and 45% of the amino acids with the respective yeast and human homologues. To substantiate this overt sequence similarity we have initiated investigation of the function and subcellular localisation of the protein product of this gene in plants.

Since no EST sequences corresponding to the predicted Atsuv3 gene were available, the predicted coding region was amplified by PCR from first strand cDNA of *A. thaliana* total RNA. The product of 1290 nucleotides generally confirms the genomic predictions, but corrects the predicted ORF in two positions at intron/exon borders. The 3′ border of the first intron in the coding region is about 120 nucleotides further downstream, and the fourth intron begins about 50 nucleotides further downstream than predicted from the genomic sequence. The previous prediction of the latter intron donor sites was probably erroneous because of the presence of a GC 5′ border instead of the more common GT bases (EMBL accession number AJ132843). Comparison with suv3 sequences from other organisms indicated that a large part of the C-terminus was missing in the predicted *A. thaliana* protein. A 3′-RACE analysis determined the precise structure of this region and identified seven additional exons in three independent 3′-RACE clones, which eliminate the predicted stop codon and extend the orf to 1713 nucleotides in 16 exons (Fig. 1). The first in-frame ATG is immediately preceded by a translational stop codon and the sequence context matches the nucleotide patterns at other translation starts in plants [24].

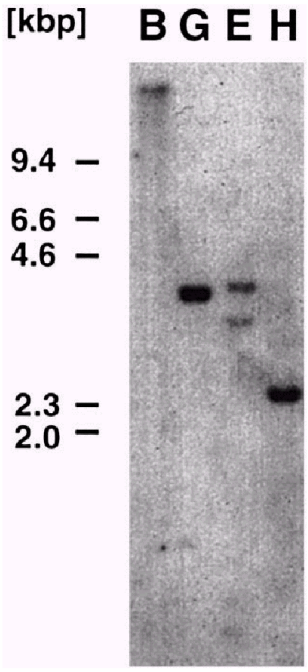


Fig. 2. The Atsuv3 gene is a unique sequence in the *A. thaliana* nuclear genome. Restriction fragments identified with the genomic Atsuv3 probe are with *Bam*HI (B) larger than 28 kb, with *Bgl*II (G) 4.114 kb, with *Eco*RI (E) 3.556, 4.251 and 0.734 kb and with *Hin*dIII (H) 2.566 kb respectively. Lengths of DNA marker fragments are given in kb.

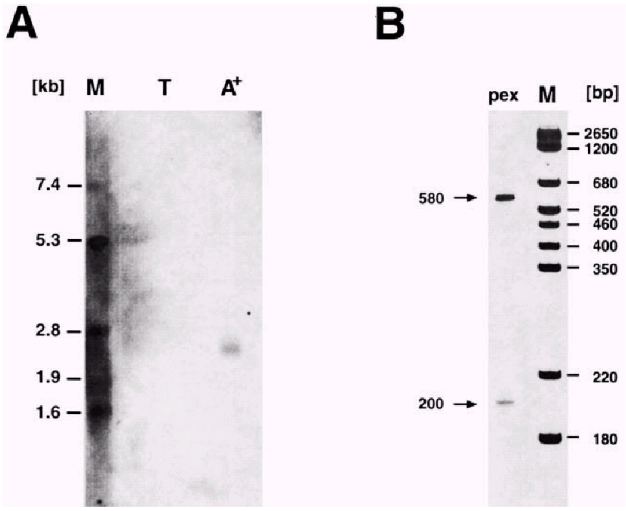


Fig. 3. The Atsuv3 gene is transcribed at a very low level in *A. thaliana*. A: Northern blot analysis using 10 µg of total RNA (lane T) and 5.7 µg of poly(A)$^+$ RNA (lane A$^+$). Sizes of RNA marker molecules are given in kb. B: Primer extension analysis of 3 µg poly(A)$^+$ RNA using a primer located 40 nucleotides 3′ to the initiation codon detects a major 5′ terminus about 540 nucleotides 5′ to the ATG. The origin of a minor signal of 200 nucleotides from the primer, 160 nucleotides upstream of the ATG, is as yet unclear.

The correspondence between the cDNA sequences and the identified genomic locus is confirmed by stringent Southern blot analysis, which identifies only a single genomic location in the total cellular DNA of *A. thaliana* (Fig. 2). The sizes of the respective restriction fragments identified in this Southern blot of digested nuclear DNA match the predictions of the genomic sequence of this region (accession number Z97337) in all instances (Fig. 2). The probe used did not detect a second suv3-like sequence (accession number AB010077) which is very divergent at the nucleotide level and is most likely a pseudogene as detailed below.

3.2. Characteristics of the deduced AtSUV3 protein

Comparison of the amino acid sequence of AtSUV3 with other SUV3 sequences from yeast, human and nematode shows that the predicted plant protein is between 76 and 215 amino acids shorter. Several stretches of amino acids are absent in both N- and C-terminal regions, while the central portion containing all of the domains typical of DExH RNA helicases is well conserved between the *A. thaliana* polypeptide and the respective proteins in various organisms, suggesting that the *A. thaliana* sequence encodes a functional polypeptide homologue (Fig. 1). Two conserved features include the signature of an ATP/GTP binding site motif

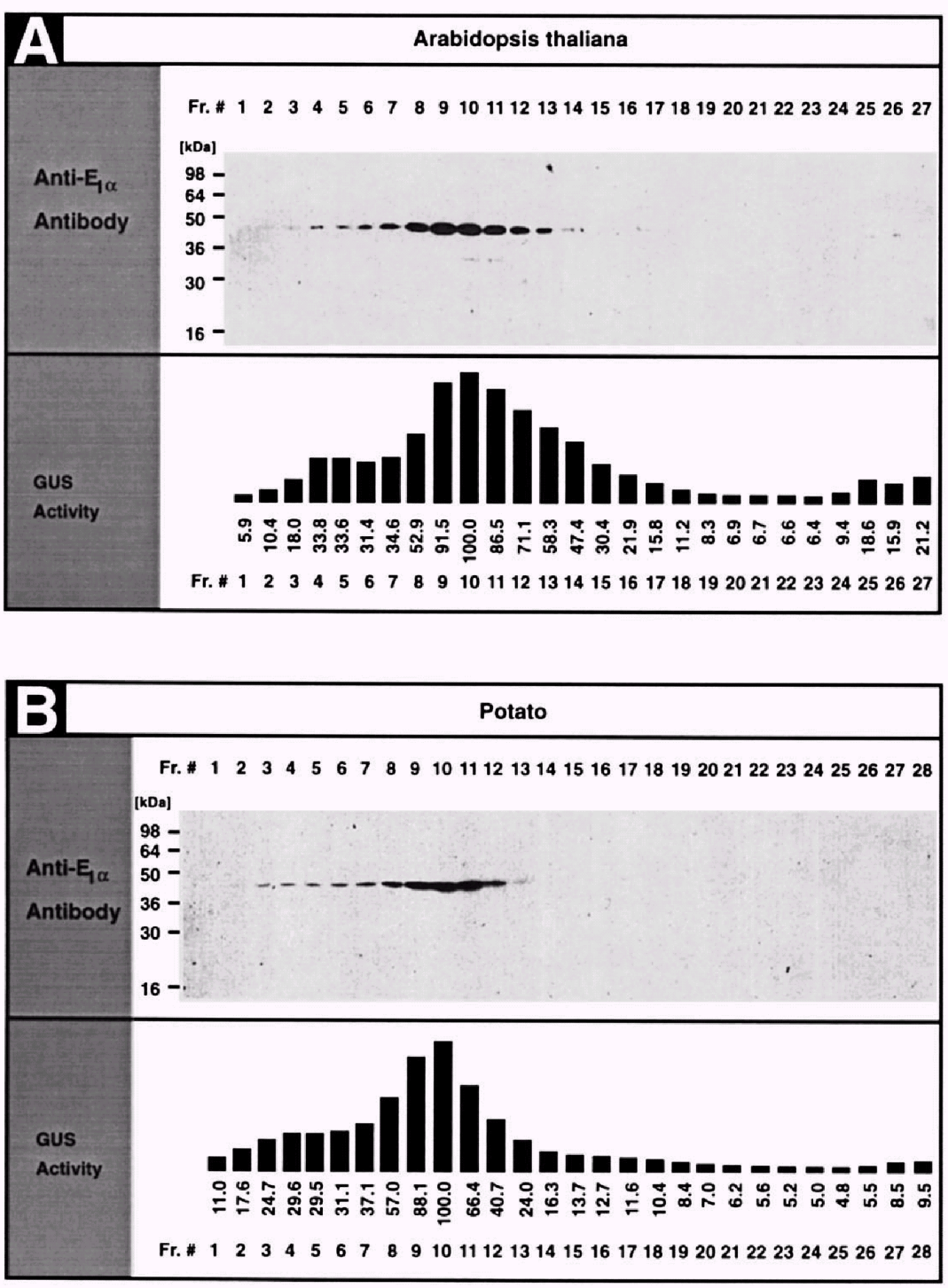


Fig. 4. The Atsuv3 gene product is a mitochondrial protein. The intracellular localisation of the Atsuv3-encoded protein was analysed in vivo in transgenic *A. thaliana* (A) and potato (B) plants, which had been transformed with the first 387 codons of the Atsuv3 open reading frame fused upstream of the uidA (GUS) reporter gene. The sedimentation of mitochondria in sucrose gradient fractions (1 ml) was monitored with antibodies against a mitochondrial marker protein, the $E_{1\alpha}$ subunit of the pyruvate dehydrogenase. Fraction 1 is at the bottom of the sucrose gradient. The distribution of GUS activity (given as per cent of the highest value) mirrors the distribution of the mitochondrial protein marker.

GPTNSGKT (amino acids 96–103 in Fig. 1) characteristic of proteins with ATPase activity and a GRAGR RNA-binding motif (amino acids 350–354 in Fig. 1). The entire protein is predicted to contain 571 amino acids starting from the first in-frame ATG with an calculated molecular weight of 63.6 kDa.

3.3. Expression of Atsuv3 is low in A. thaliana

The Atsuv3 mRNA was expected to be weakly expressed considering that no homologous sequences have been found in any of the various cDNA libraries established for *A. thaliana*, maize or rice and only a single clone in poplar (accession number AI166413). Indeed, when probing total cellular

RNA from *A. thaliana* with the Atsuv3 coding region no specific transcript signal is detected (Fig. 3A). Only upon enrichment of poly(A)$^+$ RNA and long exposure times can a single mRNA species of 2.3–2.5 kb in length be identified.

Primer extension experiments originating 40 nucleotides 3′ to the initiation codon identify a major 5′ terminus of the transcript about 540 nucleotides 5′ to the initiation codon (Fig. 3B), which suggests a size of about 2.5 kb for the mRNA in good agreement with the mRNA detected by Northern blot analysis. The origin of a minor primer extension signal about 150 nucleotides 5′ to the ATG is as yet unclear.

3.4. AtSUV3 is a mitochondrial protein

The yeast SUV3 protein has been identified as a mitochondrial protein [23]. The signal prediction program PSORT locates the plant protein deduced from the *A. thaliana* cDNA with an 85% probability to the mitochondrial matrix versus a 60% probability for a location in the nucleus (data not shown). The submitochondrial sorting to inner mitochondrial membrane and intermembrane space is predicted with 54 and 50% probabilities, respectively. The N-terminal region of the 63.6 kDa precursor protein is predicted to be cleaved after 22 amino acids resulting in a mature polypeptide of 61.2 kDa.

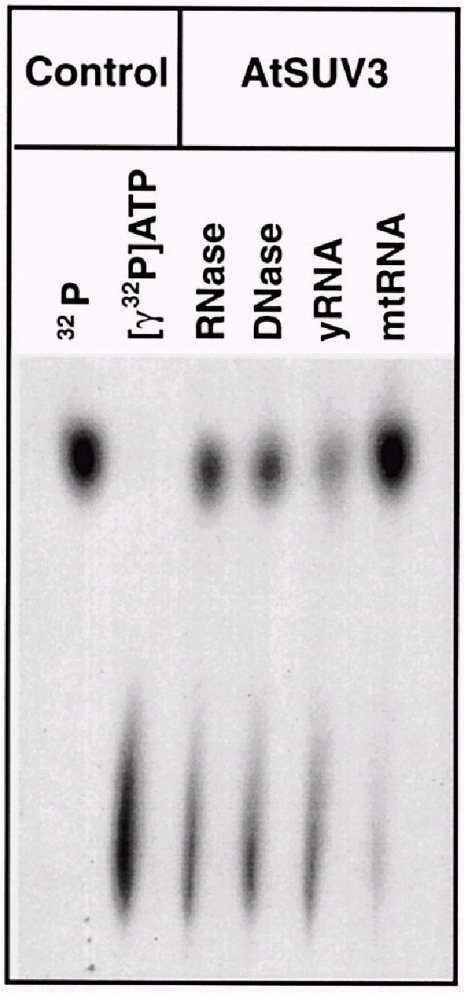


Fig. 5. ATP hydrolysis activity of the overexpressed and purified AtSUV3 protein (amino acids 11–387) is stimulated by mitochondrial RNA. Hydrolysis of [γ-^{32}P]ATP was monitored by thin-layer chromatography, which separates phosphate (lane ^{32}P) from the slower migrating ATP (lane [γ^{32}P]ATP). Addition of DNase I, RNase A, yeast tRNA or mtRNA respectively to the reaction is indicated in the figure.

The prediction of an N-terminal import signal is, although clearly positive, beyond the significant probability no reliable proof for the intracellular localisation. To investigate the intracellular localisation of the AtSUV3 protein and the specificity of the predicted N-terminal import signal experimentally, transgenic *A. thaliana* and potato plants expressing a fusion protein corresponding to the first N-terminal 387 amino acids of AtSUV3 fused to the GUS reporter protein were generated. Subcellular fractions from these transgenic plants were separated on sucrose gradients (Fig. 4). Distribution of the GUS activity was compared with the distribution of the mitochondrial marker enzyme pyruvate dehydrogenase. These analyses show that GUS activity and mitochondria cofractionate in all instances in *A. thaliana* as well as in potato, confirming the mitochondrial location of the AtSUV3 protein (Fig. 4). No GUS activity was detected in purified nuclei and chloroplasts by histochemical staining or fluorogenic assay indicating that AtSUV3-GUS is not imported into these organelles (results not shown).

3.5. The ATPase activity of the AtSUV3 protein is stimulated by mitochondrial RNA

Conservation of the major sequence elements of RNA helicases, including an ATP-binding domain, in the AtSUV3 protein suggests a function homologous to the SUV3 protein of yeast, which acts as an ATP-dependent helicase [23]. To investigate the enzymatic properties of the plant protein, a portion of AtSUV3 (amino acids 11–387) containing all the domains conserved with the SUV3 homologues in other organisms including the RNA helicase motifs was overexpressed in *E. coli* as a histidine-tagged protein. This tagged protein was purified and when tested for ATP hydrolysis activity shows a low endogenous ATPase activity (Fig. 5 and data not shown). To exclude the influence of nucleic acids potentially contaminating the recombinant protein, the ATP hydrolysis assay was repeated in the presence of RNase A or DNase I (Fig. 5). The low endogenous ATPase activity was not affected by the presence of these nucleases. While the addition of yeast tRNA had no significant effect on the reaction, clear stimulation of the ATPase activity was observed in the presence of purified mitochondrial RNA, which confirms that AtSUV3 is indeed an RNA helicase (Fig. 5).

4. Discussion

The increasing body of nucleic acid sequence data in plants generates an increasing number of sequences with similarity to genes in other organisms. Similarity alone, however, can be deceptive, functions may have changed during the course of evolution with the associated differentiation of molecular mechanisms and metabolic pathways. Thus a clear functional analysis of individual homologues must include characterisation of the subcellular localisation and an analysis of the biochemical-catalytic activities of the respective gene product. Accordingly in this report we describe the complete sequence of Atsuv3 cDNA from *A. thaliana*, which is a plant homologue of the yeast suv3 mitochondrial RNA helicase, and

present experimental evidence for the expression of the gene, the subcellular localisation of the encoded protein as well as a partial functional characterisation.

4.1. Atsuv3 is a mitochondrial protein

We conclude that the polypeptide encoded by the Atsuv3 gene is a mitochondrial protein by two independent lines of evidence. Firstly, the increasingly sophisticated signal prediction program PSORT gives the N-terminus a significant score as a mitochondrial import sequence. Secondly, experiments with two different transgenic plant species expressing the AtSUV3-GUS fusion protein clearly locate the reporter activity and thus the fusion protein in the mitochondrial compartment (Fig. 4).

4.2. Pseudogenes with similarity to suv3 in A. thaliana and maize

An incomplete open reading frame without any introns that could encode a protein with significant similarity to the yeast SUV3 protein has recently been identified in the nuclear genome of *A. thaliana* (accession number AB010077) [25]. This sequence region shows only low sequence similarity with the Atsuv3 coding region, but is more similar to two sequences found in maize genomic DNA (accession numbers X15406 and X15407). These maize loci were originally identified as pseudogenes for the GAPDH gene family and contain the Atsuv3 similarities on the opposite strand. Numerous stop codons and frame shifts mark these sequences as pseudogenes in maize. Only one of these two maize sequences (X15406) is an obvious homologue of the incomplete open reading frame in *A. thaliana* (accession number AB010077), which thus must be evolutionarily derived from a common ancestral gene, a distant relative of the genuine Atsuv3. The *A. thaliana* sequence (accession number AB010077) does contain all of the conserved central domains of SUV3 homologues, while the N-terminal region has not yet been analysed [25]. Because of the absence of any introns in the suv3-like sequence (accession number AB010077), this sequence is possibly a pseudogene derived from a reverse transcribed and integrated cDNA sequence. Expressed genes are generally interrupted by numerous introns even in the compact genome of *A. thaliana* as evidenced by the genuine Atsuv3 gene with 15 introns.

4.3. Atsuv3 shows characteristics of an RNA helicase

The presence of genuine suv3 sequences in representatives of all groups of metazoa indicates an important function of this gene in RNA metabolism in mitochondria. Intensive functional analysis of the yeast homologue to the AtSUV3 protein has shown that this mitochondrial protein is probably involved in RNA processing including 3′ maturation of mRNAs and degradation of certain group I introns [11,26]. In the mitochondrial genome of *A. thaliana* none of the genes contains a group I intron [27] and RNA 3′ processing involves different signal sequences from those found in yeast mitochondria [28], suggesting that the SUV3 protein must have other, more general function(s) in plant mitochondrial RNA metabolism. The extent of the evolutionary distance between the fungal and the plant enzymes, however, has probably not altered the enzymatic activity of the RNA helicase as such. Indeed, the plant AtSUV3 protein when expressed in *E. coli* exhibits the enzymatic hallmark of ATP-dependent RNA helicases, the stimulation of ATP hydrolysis by RNA.

Further characterisation of the AtSUV3 protein, its possible association with other proteins in a larger functional complex as well as the identification of its endogenous substrates is now feasible with this initial characterisation.

Acknowledgements: We thank Gulsen Agkun for the potato plant transformation experiments. This work was supported by a grant of the Deutsche Forschungsgemeinschaft to S.B. (Bi 590/3-3) and of the Biotechnology and Biological Sciences Research Council (G05136) to D.G. and C.J.L. C.J.L. was the recipient of a Humboldt prize during the course of this collaboration. J.K. is a fellow of the Studienstiftung des Deutschen Volkes.

References

[1] Stern, D.B. and Gruissem, W. (1987) Cell 15, 1145–1157.
[2] Dombrowski, S., Brennicke, A. and Binder, S. (1997) EMBO J. 16, 5069–5076.
[3] Hayes, R., Kudla, J., Schuster, G., Gabay, L., Maliga, P. and Gruissem, W. (1996) EMBO J. 15, 1132–1141.
[4] Kudla, J., Hayes, R. and Gruissem, W. (1996) EMBO J. 15, 7137–7146.
[5] Lisitsky, I., Klaff, P. and Schuster, G. (1996) Proc. Natl. Acad. Sci. USA 93, 13398–13403.
[6] Sarkar, N. (1997) Annu. Rev. Biochem. 66, 173–197.
[7] Gagliardi, D. and Leaver, C.J.L. (1999) EMBO J. 18, 3757–3766.
[8] Belasco, J.G. and Higgins, C.F. (1988) Gene 72, 15–23.
[9] Klaff, P. (1995) Nucleic Acids Res. 23, 4885–4892.
[10] Py, B., Higgins, C.F., Krisch, H.M. and Carpousis, A.J. (1996) Nature 381, 169–172.
[11] Margossian, S.P., Li, H., Zassenhaus, H.P. and Butow, R.A. (1996) Cell 84, 199–209.
[12] Dmochowska, A., Golik, P. and Stepien, P.P. (1995) Curr. Genet. 28, 108–112.
[13] Min, J.J., Heuertz, R.M. and Zassenhaus, H.P. (1993) J. Biol. Chem. 268, 7350–7357.
[14] Dziembowski, A., Malewicz, M., Minczuk, M., Gloik, P., Dmochowska, A. and Stepien, P.P. (1998) Mol. Gen. Genet. 260, 108–114.
[15] Bevan, M.W. et al. (1998) Nature 391, 485–488.
[16] Frohman, M.A., Dush, M.K. and Gail, R.M. (1988) Proc. Natl. Acad. Sci. USA 85, 8998–9002.
[17] Binder, S., Hatzack, F. and Brennicke, A. (1995) J. Biol. Chem. 270, 22182–22189.
[18] Bent, A.F. and Clough, S.J. (1998) in: Plant Molecular Biology Manual, 2nd edn. (Gelvin, S.B. and Schilperoort, R.A., Eds.), section B7, pp. 1–14, Kluwer Academic, Dordrecht.
[19] Sheerman, S. and Bevan, M.W. (1988) Plant Cell Rep. 7, 13–16.
[20] Leaver, C.J., Hack, E. and Forde, B.G. (1983) Methods Enzymol. 97, 476–484.
[21] Jefferson, R.A. (1987) Plant Mol. Biol. Rep. 5, 387–405.
[22] Luethy, M.H., David, N.R., Elthon, T.E., Miernyk, J.A. and Randall, D.D. (1995) J. Plant Physiol. 145, 443–449.
[23] Stepien, P.P., Margossian, S.P., Landsman, D. and Butow, R.A. (1992) Proc. Natl. Acad. Sci. USA 89, 6813–6817.
[24] Joshi, C.P., Zhou, H., Huang, X. and Chiang, V.L. (1997) Plant Mol. Biol. 35, 993–1001.
[25] Dmochowska, A., Kalita, K., Krawczyk, M., Golik, P., Mroczek, K., Lazowska, J., Stepien, P. and Bartnik, E. (1999) Acta Biochim. Polon. 46, 155–162.
[26] Conrad-Webb, H., Perlman, P.S., Zhu, H. and Butow, R.A. (1990) Nucleic Acids Res. 18, 1369–1376.
[27] Unseld, M., Marienfeld, J.R., Brandt, P. and Brennicke, A. (1997) Nature Genet. 15, 57–61.
[28] Min, J.J. and Zassenhaus, H.P. (1993) Mol. Cell. Biol. 13, 4167–4173.

6.1.2 Transcript lifetime is balanced between stabilizing stem-loop structures and degradation-promoting polyadenylation in plant mitochondria

MOLECULAR AND CELLULAR BIOLOGY, Feb. 2001, p. 731–742 Vol. 21, No. 3
0270-7306/01/$04.00+0 DOI: 10.1128/MCB.21.3.731–742.2001
Copyright © 2001, American Society for Microbiology. All Rights Reserved.

Transcript Lifetime Is Balanced between Stabilizing Stem-Loop Structures and Degradation-Promoting Polyadenylation in Plant Mitochondria

JOSEF KUHN, ULRIKE TENGLER, AND STEFAN BINDER*

Molekulare Botanik, Universität Ulm, 89069 Ulm, Germany

Received 29 August 2000/Returned for modification 26 September 2000/Accepted 24 October 2000

To determine the influence of posttranscriptional modifications on 3′ end processing and RNA stability in plant mitochondria, pea *atp9* and *Oenothera atp1* transcripts were investigated for the presence and function of 3′ nonencoded nucleotides. A 3′ rapid amplification of cDNA ends approach initiated at oligo(dT)-adapter primers finds the expected poly(A) tails predominantly attached within the second stem or downstream of the double stem-loop structures at sites of previously mapped 3′ ends. Functional studies in a pea mitochondrial in vitro processing system reveal a rapid removal of the poly(A) tails up to termini at the stem-loop structure but little if any influence on further degradation of the RNA. In contrast 3′ poly(A) tracts at RNAs without such stem-loop structures significantly promote total degradation in vitro. To determine the in vivo identity of 3′ nonencoded nucleotides more accurately, pea *atp9* transcripts were analyzed by a direct anchor primer ligation-reverse transcriptase PCR approach. This analysis identified maximally 3-nucleotide-long nonencoded extensions most frequently of adenosines combined with cytidines. Processing assays with substrates containing homopolymer stretches of different lengths showed that 10 or more adenosines accelerate RNA processivity, while 3 adenosines have no impact on RNA life span. Thus polyadenylation can generally stimulate the decay of RNAs, but processivity of degradation is almost annihilated by the stabilizing effect of the stem-loop structures. These antagonistic actions thus result in the efficient formation of 3′ processed and stable transcripts.

Control of the amount of translatable mRNA is a common parameter to regulate gene expression in many organisms. The available quantity of an RNA depends largely on the rate of synthesis and/or posttranscriptional processes which control the stability of a transcript by preventing it from degradation. Such posttranscriptional processes have been studied in prokaryotes and different subcellular compartments of eukaryotic cells. In spite of the many variations, some common features behind these processes emerge in all organisms and systems. These include the involvement of mRNA secondary structures and 3′ polyadenylation of RNA. While organized RNA secondary structures generally support transcript stabilization, polyadenylation has opposing functions in eukaryotes and prokaryotes.

In the nucleus of eukaryotic cells, pre-mRNAs are polyadenylated at 3′ ends that have been generated by endonucleolytic cleavages. These poly(A) tails play important roles in translation initiation and mRNA export from the nucleus and also have a major function in stabilizing mRNAs (27). A completely different function has been attributed to polyadenylation in prokaryotes, where the degradation of mRNA is stimulated by the presence of 3′ poly(A) tracts (4). The degradation-promoting effect of polyadenylation was even found in transcripts, which were otherwise protected and stabilized by stem-loop structures (3). A similar effect has also been observed in chloroplasts, where polyadenylation also accelerates the decay of mRNA (15, 25).

Polyadenylation has also been found in mitochondria of various organisms. The addition of adenosines to mRNAs in mammalian mitochondria has been suggested to create functional translation stop codons, which are not encoded in the DNA (22). Short oligo(A) tails were also detected at the 3′ ends of RNA in yeast mitochondria. These are about 8 nucleotides (nt) long and predominantly found at the large rRNAs (29). Similarly, oligo(A) tracts were found at some RNA species of the fragmented rRNA in *Plasmodium falciparum* (12). While the function of the oligo(A) tails in yeast is still unclear, those in *Plasmodium* were suggested to protect these RNA fragments from exonucleolytic digestion. Recently, polyadenylation was also reported for RNAs in plant mitochondria. In sunflower a dicistronic mRNA associated with cytoplasmic male sterility was suggested to be specifically destabilized by polyadenylation in fertility-restored lines (10). In maize, polyadenylation was detected at multiple sites of the *cox2* transcripts. As in sunflower, poly(A) tracts were indicated to decrease mRNA stability, although the effect observed in maize mitochondrial extracts was very weak (19). We report here nonencoded nucleotides at the 3′ ends of pea and *Oenothera* mitochondrial transcripts containing double inverted repeats in their 3′ untranslated regions (UTRs). These structures have been suggested to increase the stability of the RNA, and the stem-loop structures of pea *atp9* transcripts as well as of *orf138* mRNAs in *Brassica* were found to act as processing signals and/or stability elements at least in vitro (1, 5, 6, 16, 23). In this report special emphasis was dedicated to the interaction of the

* Corresponding author. Mailing address: Abt. Molekulare Botanik, Universität Ulm, Albert-Einstein-Allee, D-89069 Ulm, Germany. Phone: 49 731 502 2625. Fax: 49 731 502 2626. E-mail: stefan.binder@biologie.uni-ulm.de.

stabilizing effect of such double stem-loop structures and degradation-enhancing polyadenylation. In addition to adenosines, nonencoded cytidines and, to a lesser extent, uridines were detected in vivo. The combination of nonencoded cytidines with adenosines may indicate an involvement of the tRNA terminal nucleotidyltransferase in the addition of nonencoded nucleotides to 3′ ends of plant mitochondrial mRNAs.

MATERIALS AND METHODS

3′ RACE analysis of mtRNA. The 3′ rapid amplification of cDNA ends (RACE) was carried out in principle as described by Frohmann et al. (9). First-strand synthesis was initiated on 5 μg of total pea mitochondrial RNA (mtRNA) with an oligo(dT)$_{17}$-adapter primer (DTXSC) using Superscript II reverse transcriptase (RT) (GIBCO-BRL) under conditions given in the manufacturer's manual. About one-fourth of the total cDNA pool was used as the template in a PCR with adapter primer XSC and *atp9*-specific primer PA-1 (−349 to −329) with KlenTaq polymerase in a buffer supplied by the manufacturer (Clontech). After 35 cycles (30 s at 94°C, 1 min at 40°C, and 1 min at 68°C), PCR fragments were size fractionated on a 1.5% agarose gel and distinct products with sizes of 0.3, 0.4, 0.8, and 1.0 kb were observed above a smear ranging from about 0.2 to 2.0 kb. Since the PCR fragments generated from pea *atp9* transcripts, which are polyadenylated downstream of the double stem-loops, are expected to be at least 665 bp long, cDNA fragments between 0.5 and 0.9 kb in length were eluted from the gel and used as templates in a PCR with primers XSC and PA-10, a nested *atp9*-specific primer (−98 to −84 with a 5′-attached *Eco*RI restriction site). After 35 cycles (30 s at 94°C, 1 min at 50°C, and 1 min at 68°C), reaction products were digested with *Sal*I and *Eco*RI and run on an agarose gel. cDNA fragments between about 0.35 and 0.55 kb long including the expected fragment of about 0.43 kb, were eluted from the gel and cloned into pBluescript vectors (Stratagene) following standard protocols (24). In an analogous approach, oligo(dT)-adapter primer DTXSC and adapter primer XSC were replaced by DTNSX-1 and NSX-1 oligonucleotides, and *atp9*-specific primers PA-10 and PAX-1 (+259 to +274 with a 5′-attached *Xba*I restriction site) were used. cDNA clones generated with the NSX primer set carry *Nsi*I restriction sites immediately downstream of the poly(A) tracts. Linearization at this site allows the in vitro transcription of mRNAs ending with poly(A) tracts.

The 3′ RACE analysis of *Oenothera atp1* transcripts was essentially performed as described above for pea *atp9* with the following modifications. About 10 μg of *Oenothera* mtRNA and DTNSX, NSX, and *atp1*-specific primers OeatpA-RI (+1362 to +1380 with a 5′-attached *Eco*RI restriction site) and OeAXI (+1721 to +1736 with a 5′-attached *Xba*I restriction site) were used for the amplification of the 3′ regions of *atp1* transcripts. DNA fragments obtained in these PCRs were digested with *Xba*I and *Sal*I and were cloned into pBluescript II vectors. Locations of all oligonucleotides are given relative to the translation start codon.

Direct ligation of anchor primers to 3′ ends of in vitro processing products and steady-state mtRNA. In vitro processing assays were carried out under standard conditions, and reaction products were separated on a 6% polyacrylamide gel. Radioactive bands were visualized by autoradiography and were excised from the gel. The RNAs were eluted from the gel in a buffer containing 0.5 M ammonium acetate, 0.1 mM EDTA, and 0.1% sodium dodecyl sulfate. After phenol-chloroform extraction, the in vitro processing products were ethanol precipitated and ligated to 20 pmol of anchor primer (RiboNSX) in the presence of 20 U of RNA ligase (GIBCO-BRL), 1% (vol/vol) dimethyl sulfoxide, 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 10 mM dithiothreitol, 0.25 mM ATP, and 0.6 μg of bovine serum albumin per μl. The ligation reaction was performed for 24 h at 15°C in a 10-μl total reaction volume. Following two rounds of phenol-chloroform extraction, nonligated anchor primers were removed by centrifugation in Microcon 30 microfiltration units at 14,000 × *g* to a final volume of 10 μl. First-strand synthesis was primed with oligonucleotide NSX-1 using the complete ligation reaction as the template and Superscript II RT. Subsequent amplification with primers PAX-1 and NSX-1 was performed with KlenTaq polymerase (Clontech) in a buffer supplied by the manufacturer. After 35 cycles (1 min at 94°C, 1 min at 50°C, and 30 s at 68°C), PCR products were digested with *Xba*I and *Sal*I and were cloned into pBluescript vectors. Respective clones were identified in a colony hybridization using Hybond-N membranes (Amersham Pharmacia Biotech) and oligonucleotide PA-17hyb (+266 to +284) as probe, following a protocol given by the manufacturer.

For the 3′ end analysis of steady-state transcripts, the procedure described above was used with the following modifications. One microgram of total mtRNA was added to the ligation reaction mixture, and primer PA-10 was used as the *atp9*-specific primer in the amplification reaction.

Lysate preparation and in vitro processing reactions. Pea mitochondrial lysate preparation and the in vitro processing reactions were performed as described previously (6).

DNA templates and in vitro transcription. Polyadenylated pea *atp9* precursor molecules are transcribed from cDNA clone 9.17 generated in the 3′ RACE approach with primers PAX-1 and NSX-1 on in vitro-transcribed RNA obtained from clone no. 9. Plasmid 9.17 was linearized with *Nsi*I and transcribed with T7 RNA polymerase. The resulting transcript (117 nt) contains the pea *atp9* double stem-loop and 43-nt upstream and 23-nt downstream sequences. The upstream part corresponds to vector sequences from the T7 transcription start point to the *Xba*I restriction site. The downstream moieties comprise 21 adenosines at the 3′ end. Transcripts containing $(A)_3$ or $(A)_{10}$ poly(A) tracts downstream of the inverted repeat are directly transcribed from DNA fragments obtained by amplification with primer M13 universal and primers IR^+dT_3 [$(A)_3$] and IR^+dT_{10} [$(A)_{10}$] on clone 9.17 as the DNA template.

The nonpolyadenylated standard substrate with original mitochondrial upstream sequences is generated by the transcription of PCR products as previously described by Dombrowski et al. (6).

The pea *atp9* RNA substrate without inverted repeat (101 nt) is transcribed from a DNA template amplified by PCR with primers $T7IR^-$ (+159 to +179 with a T7 promoter attached at the 5′ end) and PIR^- (+241 to +258). To generate analogous substrates with $(A)_{21}$ (122 nt), $(A)_{10}$ (111 nt), and $(A)_3$ (104 nt) 3′ homopolymer sequences, respectively, primers IR^-dT, IR^-dT_{10}, and IR^-dT_3 were used instead of PIR^-.

All cloned DNA templates were inspected by complete sequence analysis. In vitro transcription reactions were performed with T7 RNA polymerase (MBI Fermentas) in a buffer supplied by the manufacturer. Purification of all in vitro transcripts by polyacrylamide gel electrophoresis was carried out as described elsewhere (6).

Nucleic acids. Pea and *Oenothera* mtRNAs were isolated from respective organelles enriched by differential centrifugation and purified on Percoll (Amersham Pharmacia Biotech) gradients as described previously (2). Sequences of the primers used are available on request. All oligonucleotides were purchased from GIBCO-BRL.

Miscellaneous methods. Sequencing reactions were carried out with the Thermo Sequenase fluorescent labeling kit (Amersham Pharmacia Biotech) according to instructions given by the manufacturer. All standard methods were used as described by Sambrook et al. (24).

RESULTS

Poly(A) tails at the 3′ ends of *atp9* mRNAs in pea mitochondria. The pea *atp9* gene is transcribed into three different mRNA species with differing 5′ termini but identical 3′ ends. These 3′ termini were mapped in two independent S1 protection analyses within the 3′ part of the second of two consecutive inverted repeats and immediately downstream thereof (6, 21). Recently it was found that this double inverted repeat with the potential to fold into a double stem-loop structure does not terminate transcription in vitro but rather functions as a processing signal and stabilizing element during a controlled 3′ processing event (5, 6). To investigate whether nonencoded poly(A) tails are present at these transcript termini, a 3′ RACE analysis, including two rounds of PCR with an oligo(dT)-adapter primer and two nested *atp9*-specific primers, was performed with total pea mtRNA (for details see Materials and Methods). The sequence analysis of 18 clones indicated the presence of nonencoded poly(A) tails at the 3′ ends of these mRNAs. In four clones, polyadenylation sites are located at positions +6 and +2 relative to the first nonpaired nucleotide (+1) downstream of the second stem (Fig. 1). The majority of the poly(A) tracts are found at positions −2 to −5 within the downstream part of the second stem. These polyadenylation sites coincide with previously mapped 3′ ends. Three clones contained poly(A) sequences attached within the second loop,

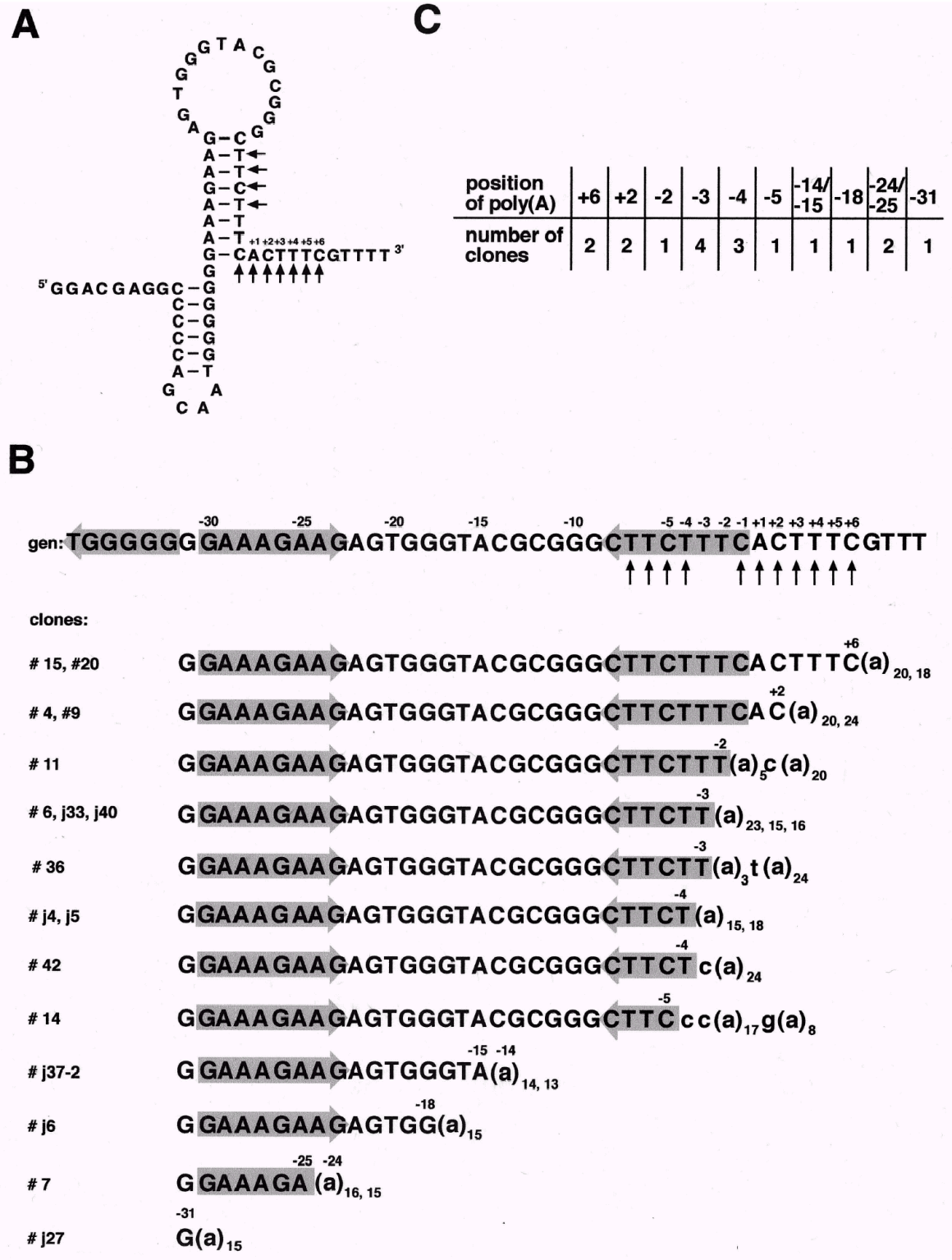


FIG. 1. 3′ RACE analysis of pea mitochondrial *atp*9 transcripts. (A) Previously mapped 3′ ends are indicated within the pea *atp*9 potential double stem-loop structure (arrows). (B) Nucleotide sequences of 3′ RACE clones containing poly(A) tails. Nonencoded nucleotides found in 18 analyzed cDNA clones (designations are shown on the left) are given in lowercase letters. The positions of the polyadenylation sites (indicated above the sequences) correspond to the 3′-most nucleotide. Numbering refers to the first nucleotide (+1) downstream of the inverted repeats (indicated by oppositely oriented grey shadowing arrows) in the genomic sequence (gen). Vertical arrows beneath the genomic sequence indicate previously mapped 3′ ends. (C) The numbers of clones identified are given for each polyadenylation site.

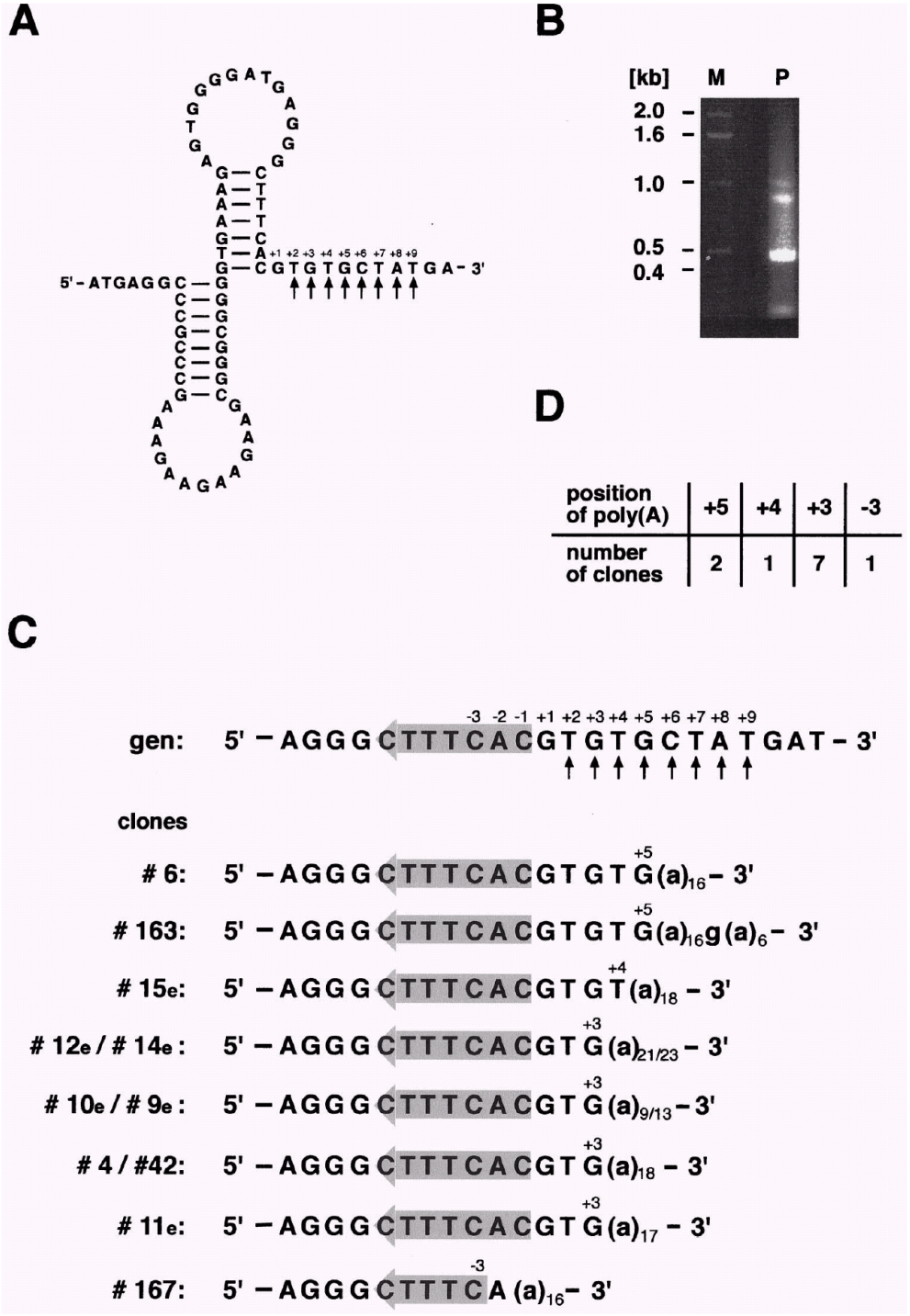


FIG. 2. Identification of potential polyadenylation sites in *Oenothera atp1* transcripts. (A) Potential double stem-loop structure found in the 3′ UTR of *Oenothera atp1* mRNAs. Previously mapped 3′ termini are indicated by arrows. (B) DNA fragments generated in a PCR with adapter primer (NSX) and an *atp1*-specific oligonucleotide (OeatpA-RI) on oligo(dT)-adapter-primed cDNA are size fractionated on a 1.5% agarose gel (lane P). Besides minor fragments of about 1.0, 0.85, and 0.2 kb, a strong cDNA fragment (0.45 kb) corresponding to RNAs polyadenylated downstream of the stem-loop is detected. Lengths of coelectrophoresed DNA marker fragments (lane M) are given in kilobase pairs on the left. (C) Partial sequences of *Oenothera atp1* cDNA clones containing nonencoded poly(A) tails (indicated in lowercase letters). Designations and numbering are analogous to those for Fig. 1. (D) Correlation of the number of clones with poly(A) addition sites.

while two others are added at the beginning and end of the upstream part of the second stem, respectively. The lengths of the poly(A) stretches vary between 14 and 25 adenosines, with 10 of the 18 clones containing more than the 17 adenosines derived from the oligo(dT)-adapter primer. In four clones, nonencoded cytidine, guanosine, and thymidine (corresponding to uridine in the RNA) nucleotides are found predominantly in the 5′ part of the attached sequences. Almost all clones investigated

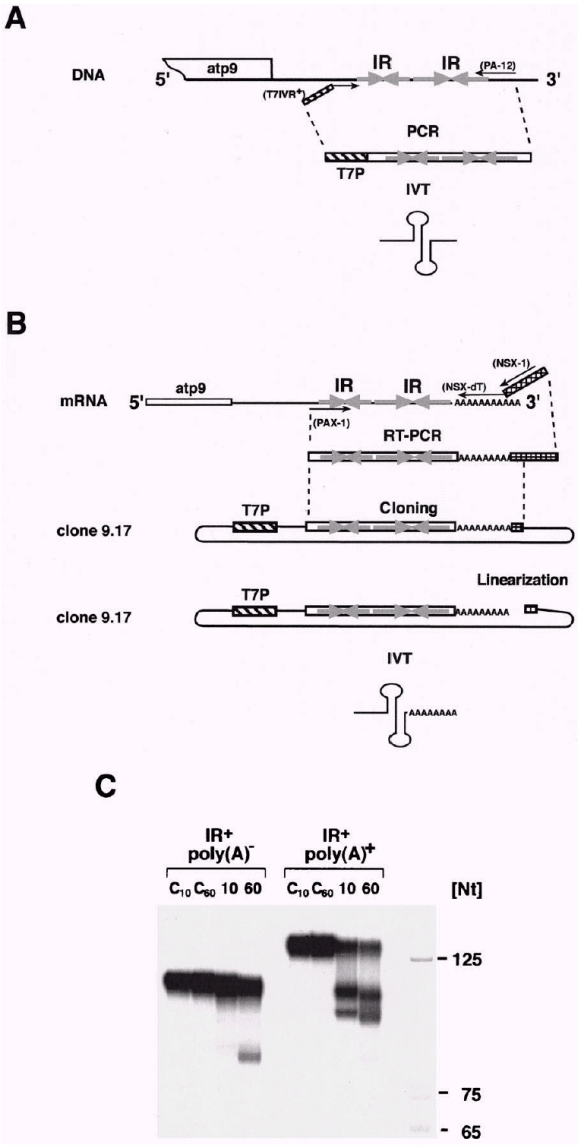


FIG. 3. In vitro processing of polyadenylated and nonpolyadenylated pea *atp9* precursor molecules. (A) Generation of the pea *atp9* standard transcript containing a double inverted repeat (grey horizontal arrows). The substrate is transcribed in vitro (IVT) from a PCR product obtained with primers PA-12 and T7IVR$^+$, which contains a promoter for T7 RNA polymerase (T7P, hatched box). (B) The polyadenylated substrate is transcribed from the linearized clone 9.17, obtained by RT-PCR cloning of a polyadenylated pea *atp9* mRNA with primers NSX-dT (adapter primer for cDNA synthesis), PAX-1, and NSX-1 (PCR). (C) Identical amounts of substrates (about 12,000 cpm) with [poly(A)$^+$] and without [poly(A)$^-$] poly(A) tracts are incubated for 10 and 60 min in a pea mitochondrial lysate. Resulting reaction products are separated on a 6% polyacrylamide gel (lanes 10 and 60). Control reactions (lanes C_{10} and C_{60}) are performed under identical conditions without protein. Lengths of coelectrophoresed DNA marker fragments are given in nucleotides (Nt) in the right margin. Discrepancies between the sizes of RNA molecules and the DNA marker fragments are due to different migration velocities of RNA and DNA molecules.

are edited at position +20 (relative to the translation start codon [+1]), while a cytidine-to-uridine alteration at position +50 is found less frequently. Both editing events change the codon identity from serine to leucine (data not shown).

Potential poly(A) tracts in *Oenothera atp1* transcripts. The analysis of *Oenothera atp1* transcripts extends this study to other double-stem-loop-containing mRNAs from mitochondria of another plant species (Fig. 2). Amplification with an *atp1*-specific primer and an adapter primer on a total cDNA pool primed with an oligo(dT)-adapter primer generated a dominant product of about 0.45 kb expected of nonencoded poly(A) sequences attached downstream of the stem-loop structure (Fig. 2B). This fragment was eluted from the gel and was either used as a template in a second PCR with a nested, *atp1*-specific primer and an adapter primer or directly cloned. Sequencing of 10 cDNA clones obtained from both amplification reactions revealed poly(A) tracts attached at positions +5, +4, and +3 immediately downstream of the double stem-loop structure. These polyadenylation sites are consistent with previously mapped 3′ ends (26). Only a single clone indicated a polyadenylation site at position −3 within the downstream part of the second stem (Fig. 2C). These results show that the distribution of the polyadenylation sites in *Oenothera atp1* mRNAs is less variable than in pea *atp9* transcripts. The lengths of the poly(A) tails observed in the *Oenothera atp1* cDNA clones range from 16 to 23 adenosines with only a single guanosine insertion within a poly(A) sequence.

Rapid and efficient processing of polyadenylated pea *atp9* transcripts in vitro. The results of the 3′ RACE analyses identified nonencoded poly(A) tracts at the 3′ ends of transcripts containing double stem-loop structures in their 3′ UTRs. Such poly(A) tails were recently suggested to decrease mRNA stability in plant mitochondria (10, 19). This destabilizing effect would counteract the function of the double stem-loop, which was found to act as a processing signal and stability element at least in vitro (5, 6). To solve this question, the consequence of poly(A) tails on double-stem-loop-containing mRNA was tested in vitro. Nonpolyadenylated synthetic transcripts (96 nt) used in these assays were transcribed from a T7 promoter-containing PCR product (Fig. 3A) and corresponded to the standard substrate shown to be correctly processed in a pea mitochondrial in vitro system (6). It contains the pea *atp9* double stem-loop structure and 21 and 28 original mitochondrial nt upstream and downstream, respectively. In the polyadenylated mRNA (113 nt), the 3′ trailer downstream of the stem-loop structure consists of a 21-mer adenosine homopolymer downstream of two mitochondrially encoded nucleotides while the upstream leader (42 nt) represents pBluescript vector sequences. Details of the construction of the DNA templates for transcription of these substrates (Fig. 3B) are given in Materials and Methods. To test the influence of the regions upstream of the inverted repeat, an analogous RNA substrate, which contains original mitochondrial downstream sequences (28 nt) and the alien sequences deriving from the vector (42 nt), was also tested in vitro. After incubation of the nonpolyadenylated standard transcript for 60 min, a single processing product of about 70 nt was observed (Fig. 3C, left). This product corresponds to an RNA species where the downstream part is removed, as seen in previous processing experiments (6). An accelerated generation of processing products is observed upon incubation of polyadenylated substrates. After 10 min, two processing products of about 70 and 80 nt are generated. While the smaller product has a size consistent with a 3′ end similar to that of the product obtained

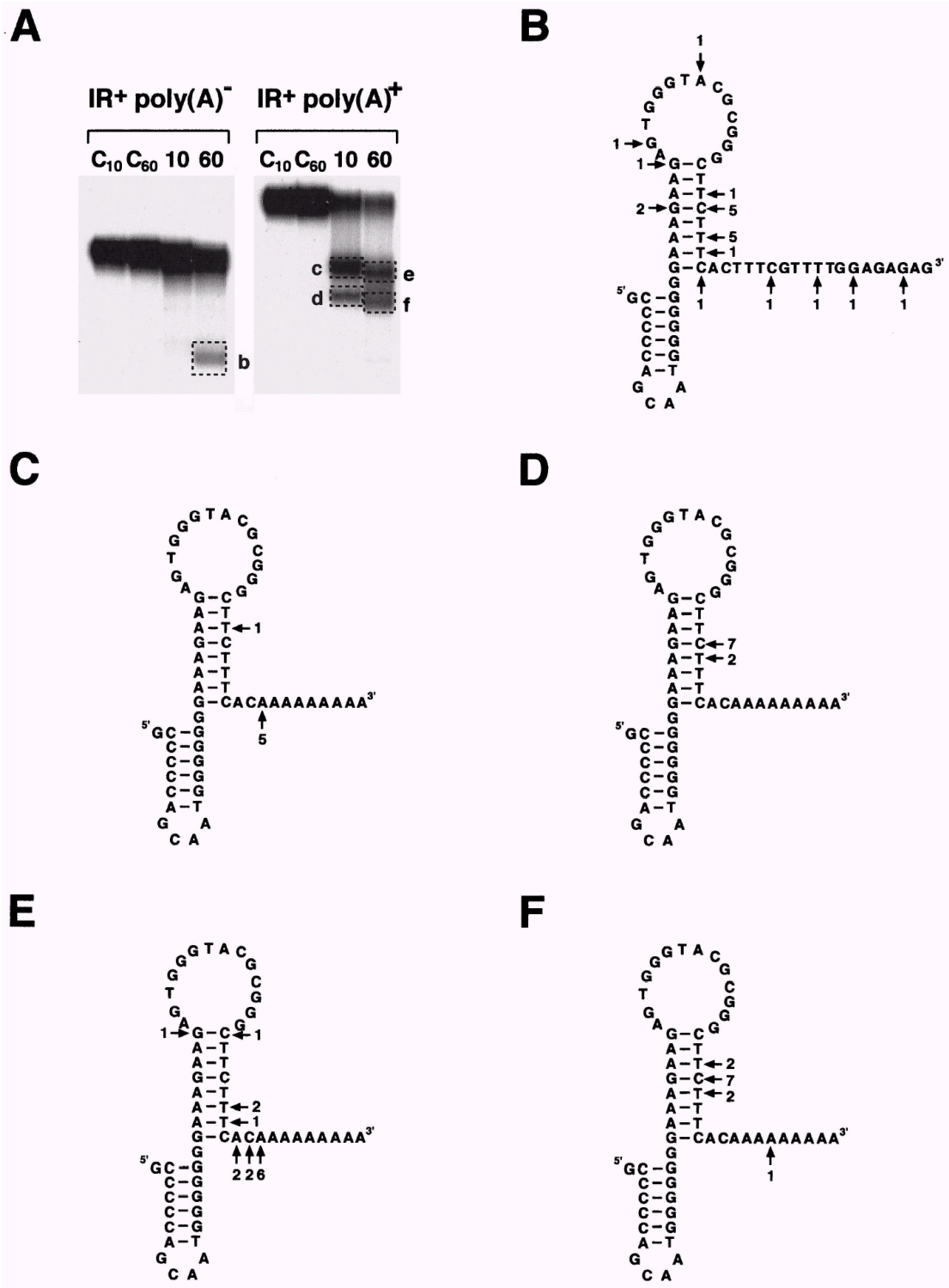


FIG. 4. Precise determination of 3′ termini of in vitro processing products. (A) The in vitro processing products (b to f, indicated by dotted boxes) obtained from nonpolyadenylated [poly(A)$^-$] and polyadenylated [poly(A)$^+$] substrates with inverted repeats (IR$^+$) were excised, and the RNA molecules were eluted from the gel. The 3′ ends of the different products were analyzed by direct anchor primer ligation followed by RT-PCR analysis. Lanes 10 and 60 represent results at 10 and 60 min; lanes C_{10} and C_{60} represent control reactions at 10 and 60 min. (B to F) Secondary structures of the pea *atp9* double stem-loop with nonpolyadenylated (B) and polyadenylated (C to F) downstream moieties. Arrows indicate the 3′ terminal nucleotide found in individual clones. The numbers of cDNA clones containing certain 3′ ends are given at the arrows. Designations of the individual stem-loops (B to F) correspond to the designations of the analyzed processing products (b to f).

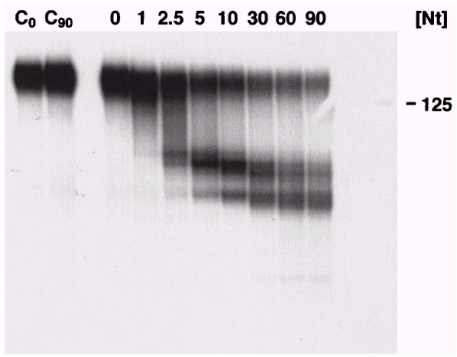


FIG. 5. Time course of the in vitro processing of a polyadenylated pea *atp9* precursor. A polyadenylated pea *atp9* substrate is incubated over a period of 90 min, and probes are taken at 0 to 90 min, as indicated above the lines. Control reactions are performed under identical conditions in the absence of protein for 0 (C_0) and 90 (C_{90}) min. Sizes of coelectrophoresed DNA marker fragments are given in nucleotides (Nt).

in the reaction with the nonpolyadenylated standard substrate, an additional RNA, about 5 to 10 nt larger, is generated. Longer incubation (for 60 min) resulted in the generation of two products that seem to be slightly smaller than those observed after 10 min (Fig. 3C, right). A product very similar to the one generated from the nonpolyadenylated standard substrate and to the smaller one (70 nt) deriving from the polyadenylated substrate was observed in in vitro tests with the nonpolyadenylated substrate, with vector sequences upstream of the stem-loop structure. The temporal dynamic of the product formation is identical to that of the processing reaction of the nonpolyadenylated standard substrate, which indicates that sequences upstream of the stem-loop structure have no degradation-promoting function (data not shown).

A similar processing pattern was observed with heterologous synthetic polyadenylated and nonpolyadenylated *Oenothera atp1* transcripts. Again, accelerated processing is observed with poly(A)$^+$ substrates, while final degradation of this RNA seems to be unaffected (data not shown).

The precise 3′ ends of the detected products were determined by a direct ligation of anchor primers (RiboNSX) to eluted RNA species, followed by an RT-PCR analysis initiated with primer NSX complementary to the anchor primer (Fig. 4). Sequencing of 22 cDNA clones obtained from the processing products generated from the nonpolyadenylated standard substrate revealed 3′ ends which scatter over a broad range (from nucleotide positions −26 to +17) with reference to the first nucleotide (+1) downstream of the second stem-loop (Fig. 4B). The 3′ ends detected in 14 clones, however, are located within the downstream part of the second stem or immediately downstream of this structure and coincide with previously mapped transcript termini (6) (Fig. 1 and 4B). In the larger gel fragment excised, other termini are identified at low frequencies, which are attributed to the background of random termini in the RNA population of such an in vitro processing reaction (Fig. 4A). Transcript termini from the smaller processing products generated after 10 and 60 min from polyadenylated substrates are detected in 9 and 12 clones, respectively, and are almost exclusively located within the second stem, which is consistent with the majority of termini generated from the

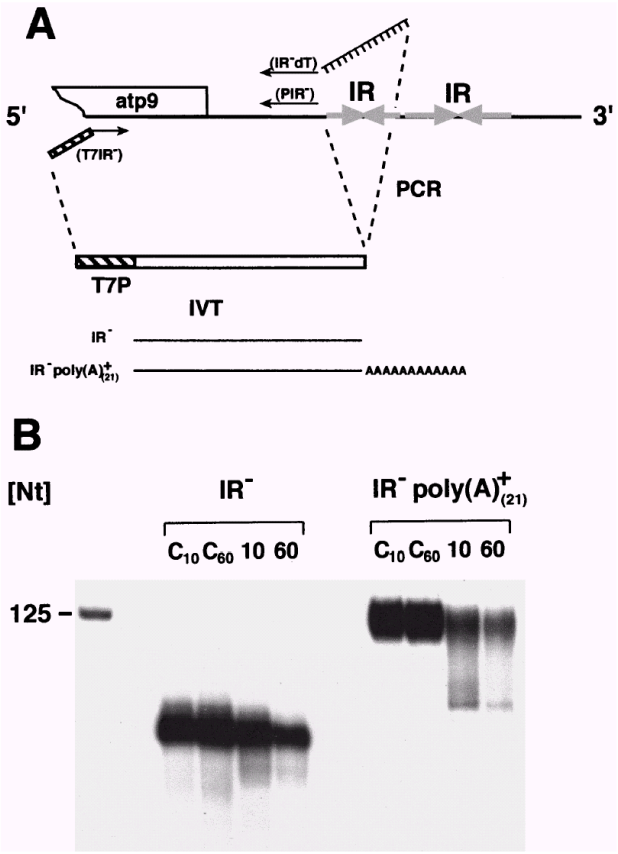
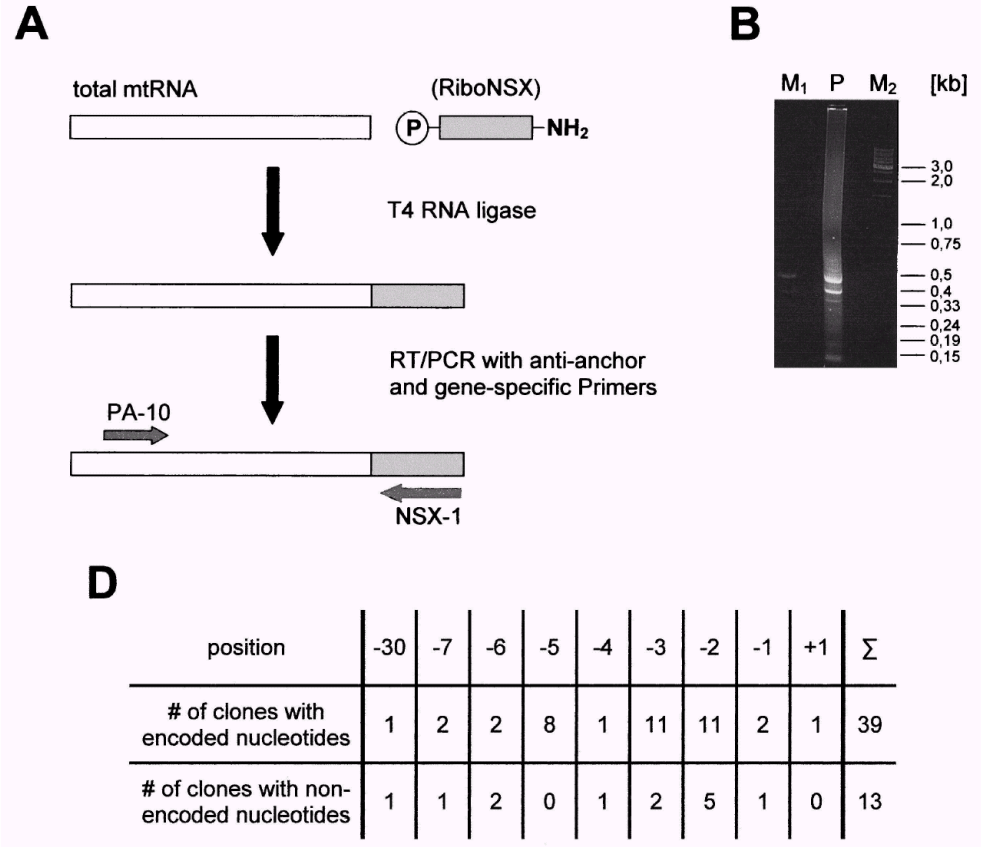


FIG. 6. Accelerated degradation of RNAs without stem-loops containing 3′ poly(A) tails. (A) Substrates with no inverted repeat containing no 3′ homopolymer tract (IR^-) and a 3′ poly$(A)_{21}$ [IR^- poly$(A)^+$] are transcribed in vitro (IVT) from PCR products with primer $T7IR^-$ containing a T7 promoter (T7P, hatched box) and primers PIR^- and IR^-dT $(A)_{21}$, respectively. (B) Incubation of the respective substrates with pea mitochondrial protein extracts for 10 and 60 min. Designation of the lanes (10 and 60) corresponds to the length of incubation. Control reactions are performed in the absence of protein under otherwise identical conditions for 10 (C_{10}) and 60 (C_{60}) min. The sizes of DNA marker fragments are indicated in nucleotides (Nt).

nonpolyadenylated precursor RNA (Fig. 4D and F). In contrast the bulk of the 3′ ends of the larger processing products are found immediately downstream of the second stem with slightly shorter products observed after 60 min (Fig. 4C and E). Interestingly, half of the 22 clones obtained from these products end with a nonencoded adenosine, indicating that the poly(A) tail is not completely removed. In summary the results of the in vitro processing assays indicate that polyadenylated transcripts are processed much faster than nonpolyadenylated substrates. Total degradation of the double-stem-loop-containing substrates, however, is not significantly accelerated.

Temporal dynamics of the in vitro processing of polyadenylated pea *atp9* transcripts. The generation of an additional distinct product from polyadenylated transcripts raises the possibility that this could be an intermediate processing product. To investigate the temporal course of the in vitro processing of polyadenylated pea *atp9* mRNAs (Fig. 1B), the generation of processing products was followed over a period of 90 min (Fig.



D

position	-30	-7	-6	-5	-4	-3	-2	-1	+1	Σ
# of clones with encoded nucleotides	1	2	2	8	1	11	11	2	1	39
# of clones with non-encoded nucleotides	1	1	2	0	1	2	5	1	0	13

FIG. 7. A direct anchor primer ligation–RT-PCR analysis identifies nonencoded nucleotides at pea *atp9* steady-state transcripts. (A) Scheme of the direct anchor primer ligation–RT-PCR approach as outlined in Materials and Methods. (B) Ethidium bromide-stained agarose gel with DNA fragments obtained from a PCR carried out with primers PA-10 and NSX-1. Sizes of DNA marker fragments are given in kilobase pairs on the right. (C) Sequences of the downstream inverted repeat of the pea *atp9* 3′ UTR (indicated by grey shading) obtained from a direct anchor primer ligation–RT-PCR approach. Nonencoded nucleotides are given in lowercase letters. Previously mapped 3′ transcript termini are indicated by vertical arrows beneath the genomic sequence (gen). The numbering of the sequence is given with respect to the first nucleotide (+1) downstream of the inverted repeat. (D) Number of clones containing encoded 3′ terminal nucleotides and clones with nonencoded nucleotides found at certain positions relative to the first nucleotide downstream of the inverted repeat (+1).

5). Already after 1 min, diffuse smaller RNAs appear, indicating the initiation of the processing reaction, which results in the detectable formation of the approximately 90- to 95-nt-long RNA fragment after 2.5 min. The smeared RNA between the substrate and this product indicates exonucleolytic processivity. This impression is substantiated by the continuously decreasing size of this 90- to 95-nt-long RNA, with several nucleotides being consecutively removed upon prolonged incubation. A similar effect is also seen with the 80- to 85-nt-long product, which first appears after 5 min. The substrate and the two processing products show distinct dynamics in their respective abundances. As expected, the substrate decreases continuously over 90 min. The larger product (90 to 95 nt) has its highest abundance after 10 min, while the smaller product (80 to 85 nt) reaches maximum abundance after 60 min. These results suggest that the large processing product, whose 3′ ends map immediately downstream of the double stem-loop, is a stabilized intermediate in the exonucleolytic generation of the small product, with its 3′ ends located within the downstream part of the second stem. Thus, an exoribonuclease seems to remove the poly(A) tails very rapidly but stalls or falls away from the RNA either at the 3′ end of the stem-loop structure or simply at the transition from the poly(A) stretch to the "normal" sequence. The smaller product most likely results from a second steric hindrance within the downstream part of the stem.

Accelerated total degradation of RNA substrates with no inverted repeat. The results presented above indicate that polyadenylation accelerates RNA processing of stem-loop-containing RNAs, while the total degradation of the transcripts remains almost unaffected. To investigate the stabilizing function of this stem-loop structure during the formation of the different processing products, polyadenylated and nonpolyadenylated RNAs without a stem-loop were tested in the in vitro processing system. The substrate without a stem-loop structure and poly(A) tract, representing 101-nt-long RNA molecules with 3′ ends just upstream of the first inverted repeat (Fig. 6A), is consecutively degraded without formation of any stable intermediate, indicating that the stem-loop is indeed responsible for the generation of stabilized products (Fig. 6B, left). A different processivity is seen with polyadenylated transcripts. Here the presence of a poly(A) tract in the substrate signifi-

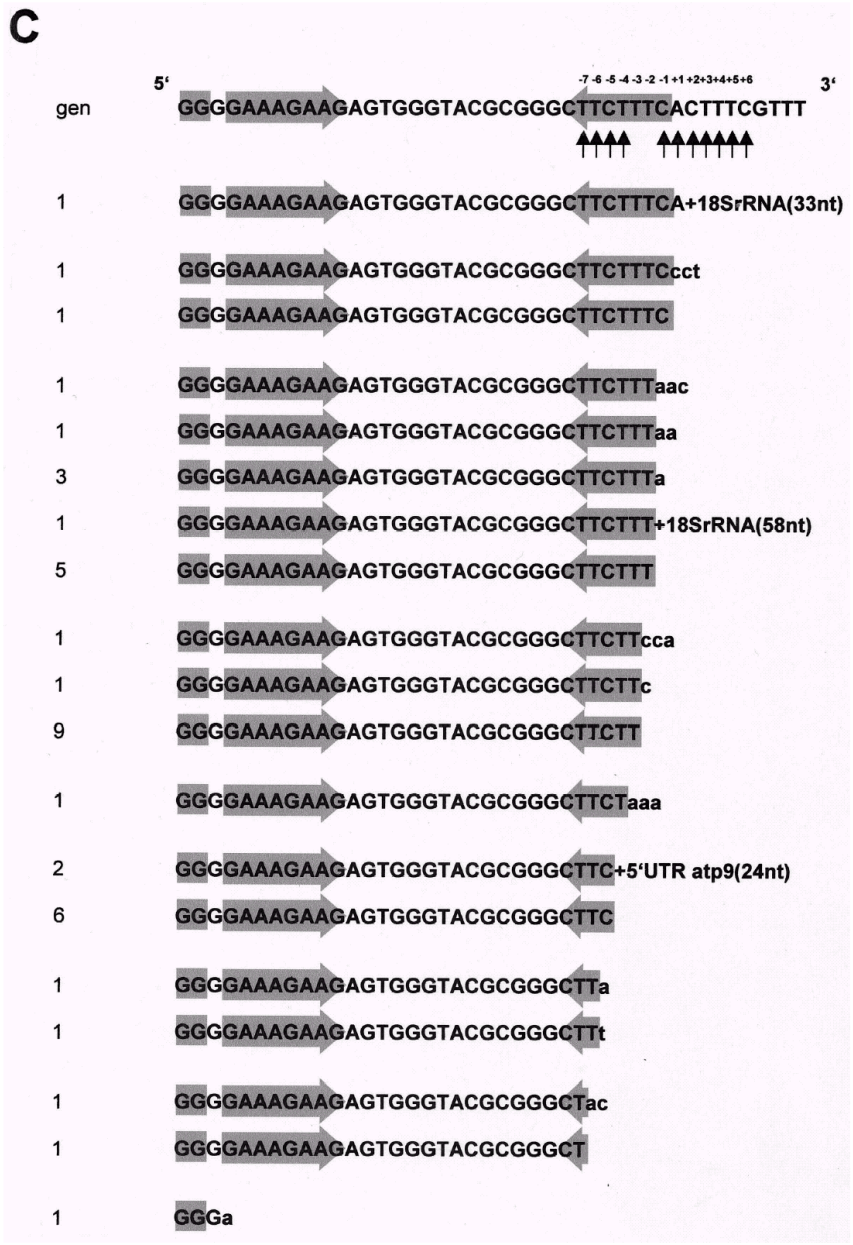


FIG. 7—*Continued.*

cantly accelerates the total turnover of the precursor RNA (Fig. 6B, center). Only a weak signal indicates the presence of an intermediate, which corresponds in size to the RNA without a poly(A) tail. The low intensity of this signal indicates only weak stability of this product.

Since the diffuse RNAs between the substrate and this intermediate are suggestive of an exonucleolytic digestion of the substrates, a potential involvement of a 3′-to-5′ exoribonuclease activity was monitored with 5′- and 3′-labeled transcripts containing the inverted repeat (Fig. 3A). While a stable product of the expected size is seen with the 5′-labeled substrates, the 3′-labeled transcripts disappear without formation of a detectable processing product or intermediate (data not shown). This suggests that indeed a 3′-to-5′ exoribonuclease takes part in the 3′ processing and/or degradation event investigated here.

Detection of nonencoded nucleotides at 3′ transcript termini of steady-state RNA. The 3′ RACE analysis using an oligo(dT)-adapter primer does indicate the presence of poly(A) sequences at the 3′ ends of mtRNAs. Several questions are, however, left unanswered in this approach. First, the lengths of the poly(A) tails cannot be determined, and second, the majority of other nonencoded nucleotides (i.e., C, G, or U) might remain undetected. This may be a major drawback, since, for example, considerable uridylyltransferase activity has been detected in plant mitochondrial protein extracts (S. Binder, unpublished results).

For an alternative approach, anchor primers were ligated directly to 3′ ends of total steady-state mtRNA from pea. The products of an RT-PCR initiated at an oligonucleotide complementary to the anchor primer were cloned and analyzed (Fig. 7A and B).

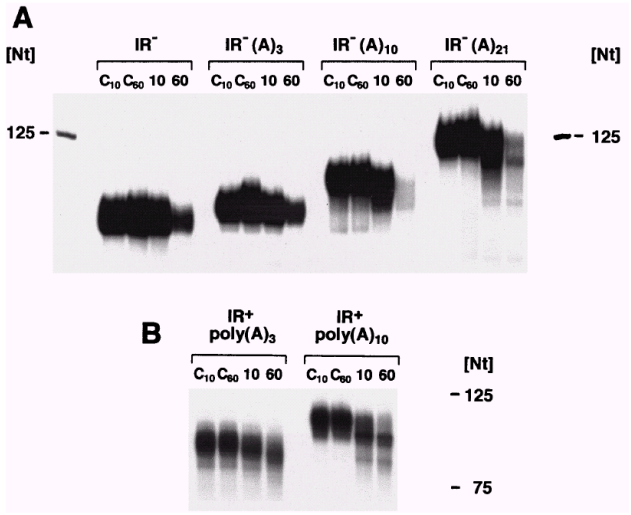


FIG. 8. Oligo(A) tails with 10 or more adenosines promote RNA degradation or processing. (A) Incubation of synthetic transcripts with no inverted repeat with either no extension (IR^-), a 3-nt extension [$IR^-(A)_3$], a 10-nt tract [$IR^-(A)_{10}$], or a 21-nt tail [$IR^-(A)_{21}$] are incubated with a pea mitochondrial lysate for 10 and 60 min. Clear differences are observed after 60 min with strong degradation-stimulating effects for the tails with 10 and 21 adenosines. In comparison, no significant differences in the decay are observed in reactions with substrates without extensions or with three additional terminal adenosines. Nt, nucleotides. (B) Substrates containing the pea *atp9* inverted repeat with 3 [$IR^+poly(A)_3$] or 10 [$IR^+poly(A)_{10}$] adenosines attached 2 nt downstream of the second stem were incubated with a pea mitochondrial lysate. The product generated from the latter substrate where sequences downstream of the stem-loops have been removed clearly indicates a processing-promoting effect of the 10-adenosine extension. In comparison, only a minor shortening of the substrate with three terminal adenosines is observed after 10 and 60 min (lanes 10 and 60). No degradation or processing is observed after incubation of the substrates for 10 and 60 min in the absence of protein (lanes C_{10} and C_{60}). Nt, nucleotides.

From 39 clones sequenced, 13 contain 3′ nonencoded nucleotides at sites previously mapped as 3′ termini (Fig. 7C and D). Adenosines were found in 10 clones, but in addition, cytidines are detected in 5 clones and thymidines (corresponding to uridines) are detected in 2 clones. The cytidines are almost exclusively combined with adenosines. Sole adenosines are found in seven clones, albeit five of these clones contain only single nonencoded adenosines. This experiment confirms that cytidines and, less frequently, uridines are added in vivo to the 3′ ends of plant mitochondrial mRNAs. RNA fragments deriving from the 18S rRNA and the 5′ UTR of the *atp9* transcript are most likely coincidentally ligated to the 3′ ends of *atp9* mRNA, although identical *atp9* 5′ UTR fragments are found in two clones derived from independent ligation–RT-PCR approaches.

Poly(A) tracts with 10 or more adenosines stimulate RNA processing and degradation. The anchor primer ligation–RT-PCR analysis detects only short nonencoded extensions up to 3 nt. Two explanations are possible for such relatively short tracts. Either the extensions per se are indeed only a few nucleotides long, or they are the result of the processing reaction. To test the latter assumption, substrates with no inverted repeat containing 0, 3, 10, or 21 additional adenosines were incubated with the pea mitochondrial lysate. While both 10 or 21 oligo(A) tails significantly accelerate degradation, no difference is observed in the processing of substrates with 3 or 0 additional adenosines (Fig. 8A). An analogous effect is seen with substrates containing inverted repeats. An extension of 3 adenosines has no impact on processing, whereas a processing product is rapidly formed from a precursor with 10 adenosines (Fig. 8B).

These results and the fairly fragment in vitro processing products containing one or more residual nonencoded adenosines (Fig. 4C to F) strongly suggest that the short extensions detected in vivo are the product of RNA processing and/or degradation.

DISCUSSION

3′ nonencoded nucleotides in plant mitochondrial transcripts. Two independent approaches were used to identify nonencoded nucleotides at the 3′ ends of plant mitochondrial transcripts. While the 3′ RACE analysis initiated with an oligo(dT)-adapter primer indirectly shows the presence of nonencoded adenosines at the 3′ transcript termini, the direct anchor primer ligation–RT-PCR approach provides straightforward evidence for the existence of such nonencoded 3′ extensions without any inherent experimental bias for certain nucleotides.

The 3′ RACE analysis predominantly detects poly(A) tracts within the 3′ part of the second stem or immediately downstream thereof. Since no adenosines are encoded in these regions downstream of both pea *atp9* and *Oenothera atp1* genes, all these clones should derive from annealing of the oligo(dT)-adapter primer at nonencoded, posttranscriptionally added adenosines (Fig. 1 and 2). Only the 3′ ends observed in pea *atp9* clones no. 7 and j27 are located in regions which would allow an annealing of the oligo(dT) stretch to encoded sequences (Fig. 1). Besides the adenosines expected from the use of an oligo(dT)-adapter primer, cytidines, thymidines (corresponding to uridines in the RNA), and guanosines are also found in the 3′ RACE products of pea *atp9* and *Oenothera atp1* mRNAs. While the location of the guanosine within the 17-mer adenosine (corresponding to thymidines) stretch in the adapter primer rather suggests an artifact, several cytidines and a single thymidine are either found directly at or near the 3′-most nucleotides. The direct anchor primer ligation–RT-PCR approach confirms the presence of cytidines and uridines in more than half of the clones with nonencoded nucleotides. Interestingly, no guanosine is found in the clones obtained in this analysis. The detection of other RNA fragments at the 3′ ends of *atp9* mRNAs is due to the high content of usually 18S rRNA fragments in mtRNA preparations. Such fragments are also often observed in chimeric clones of mitochondrial cDNA libraries (11). The validity of the method is indirectly supported by the analogous investigation of the in vitro processing products (Fig. 4B), where in 64 clones no nucleotides were detected that could be derived from artifacts generated in the ligation or amplification reactions.

The real in vivo length of the nonencoded extensions remains unclear. A diagnostic 3′ RACE analysis initiated by an oligo(dT)-adapter primer with 17 thymidines on an RNA with a 24-mer adenosine poly(A) tail results in clones containing

between 15 and 24 adenosines, corroborating the length variation occurring under such experimental conditions (data not shown). Thus, such a 3′ RACE analysis does not allow any conclusion on the a priori lengths of the poly(A) tracts. In the alternative approach, the number of nonencoded nucleotides found is rather low. The majority of these clones contain a single nonencoded nucleotide, and the maximum of three posttranscriptionally added nucleotides is found in only three clones. No longer poly(A) tails are observed in this approach, suggesting a very low frequency of such longer extensions.

The in vitro assays of substrates with poly(A) tracts of different sizes indicate that tails with 10 nucleotides already promote a similar acceleration of degradation, as do longer poly(A) tails (Fig. 8A). In contrast, short tails with only three adenosines have no influence on transcript stability, suggesting that the termini seen in the direct anchor primer ligation–RT-PCR analysis derive from relatively stable end products rather than from substrates promoting and initiating degradation. This assumption is substantiated by the detection of one or more residual nonencoded adenosines at the processing products generated from polyadenylated substrates.

The observation that mitochondria can discriminate between transcripts with homopolymer tracts of different sizes and specifically degrade those with 10 or more nucleotides supports the functional and biological significance of these nonencoded nucleotides.

Function of polyadenylation in plant mitochondria. Recent reports of polyadenylation of plant mitochondrial transcripts suggested that the presence of 3′ poly(A) tails accelerates the degradation of mRNAs (10, 19). Our results obtained for transcripts without a stem-loop structure are consistent with these results. A significant acceleration of RNA decay is caused by the presence of poly(A) tails in these transcripts. Surprisingly, an even stronger destabilizing effect is observed with poly$(G)_{21}$ tails, which contradicts previous reports on chloroplasts and cannot be rationally explained at present (7, 8; data not shown). No effects of poly(G) stretches were detected in transcripts of nuclear reporter genes in plants, indicating that indeed the effects of poly(G) stretches might differ for different compartments of the plant cell (13).

In plant mitochondria we furthermore observe a differentiated response to polyadenylation of mRNAs with double stem-loop structures. With these substrates a significant acceleration of 3′ end processing is observed, while total degradation is only marginally increased (Fig. 3). In this reaction RNA moieties downstream of the stem-loop are rapidly removed, but the progress of degradation is stopped or stalled at the stem-loop structure (Fig. 4). The stabilizing effect of these structures then seems to be stronger than the stimulation of mRNA decay by polyadenylation. The identification of poly(A) addition sites within the stem-loop indicates that these are polyadenylated again and suggests repeated cycles of polyadenylation and removal of poly(A) tails together with several adjacent nucleotides. This may be an effective mechanism for eventually overcoming the stem-loop structure and initiating the total degradation of the RNA. The life span of such a transcript is thus balanced between the degradation-promoting polyadenylation and the stabilizing effect of the stem-loop. In plant mitochondria the former effect seems to be minor in comparison to that in bacteria and chloroplasts, where poly(A) tails have a much stronger destabilizing effect on stem-loop-containing transcripts (3, 17, 18).

How are the 3′ nonencoded nucleotides added to mRNAs? The presence of nonencoded nucleotides at the 3′ ends of mRNAs raises the question of how these nucleotides are added. In other compartments of plant cells and in other organisms, poly(A) tails are generally synthesized by poly(A) polymerase, which has so far not been reported in plant mitochondria. It has recently been shown that this protein from *Escherichia coli* incorporates all 4 nt into the 3′ extensions, at least in vitro, so one can speculate that an analogous enzyme activity may add the nonencoded nucleotides to mRNA 3′ ends in plant mitochondria (28).

The presence of a significant number of cytidines among the nonencoded nucleotides also raises the possibility that a tRNA terminal nucleotidyltransferase is involved in the addition of the nonencoded nucleotides. This theory is strengthened by the observation that the cytidines are almost exclusively combined with adenosines and that complete CCA additions were identified in two clones (Fig. 1 and 7). This enzyme activity has been detected in plant mitochondrial extracts from potato and wheat and should now be tested for the potential to add short or even longer oligo(A) tracts to mRNA substrates (14, 20).

ACKNOWLEDGMENTS

This work was supported by grant Bi 590/3-3 from the Deutsche Forschungsgemeinschaft and a fellowship from the Studienstiftung des Deutschen Volkes to J.K.

We thank Axel Brennicke for fruitful discussions and his constructive comments on the manuscript.

REFERENCES

1. **Bellaoui, M., G. Pelletier, and F. Budar.** 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.
2. **Binder, S., and A. Brennicke.** 1993. Transcription initiation sites in mitochondria of *Oenothera berteriana*. J. Biol. Chem. **268:**7849–7855.
3. **Blum, E., A. J. Carpousis, and C. F. Higgins.** 1999. Polyadenylation promotes degradation of 3′-structured RNA by the *Escherichia coli* mRNA degradosome *in vitro*. J. Biol. Chem. **274:**4009–4016.
4. **Carpousis, A. J., N. F. Vanzo, and L. C. Raynal.** 1999. mRNA degradation. A tale of poly(A) and multiprotein machines. Trends Genet. **15:**24–28.
5. **Dombrowski, S., and S. Binder.** 1998. 3′-inverted repeats in plant mitochondrial mRNAs act as processing and stabilizing elements but do not terminate transcription, p. 9–12. *In* I. M. Moller, P. Gardeström, K. Glimelius, and E. Glaser (ed.), Plant mitochondria: from gene to function. Backhuys Publishers, Leiden, The Netherlands.
6. **Dombrowski, S., A. Brennicke, and S. Binder.** 1997. 3′-inverted repeats in plant mitochondrial mRNAs are processing signals rather than transcription terminators. EMBO J. **16:**5069–5076.
7. **Drager, R. G., M. Zeidler, C. L. Simpson and D. B. Stern.** 1996. A chloroplast transcript lacking the 3′ inverted repeat is degraded by 3′ to 5′ exoribonuclease activity. RNA **2:**652–663.
8. **Drager, R. G., J. Girard-Bascou, Y. Choquet, K. L. Kindle, and D. B. Stern.** 1998. *In vivo* evidence for 5′ to 3′ exoribonuclease degradation of unstable chloroplast mRNA. Plant J. **13:**85–96.
9. **Frohman, M. A., M. K. Dush, and G. R. Martin.** 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA **85:**8998–9002.
10. **Gagliardi, D., and C. J. Leaver.** 1999. Polyadenylation accelerates the degradation of the mitochondrial mRNA associated with cytoplasmic male sterility in sunflower. EMBO J. **18:**3757–3766.
11. **Giegé, P., and A. Brennicke.** 1999. RNA editing in *Arabidopsis* mitochondria effects 441 C to U changes in ORFs. Proc. Natl. Acad. Sci. USA **96:**15324–15329.
12. **Gillespie, D. E., N. A. Salazar, D. H. Rehkopf, and J. E. Feagin.** 1999. The fragmented mitochondrial ribosomal RNAs of *Plasmodium falciparum* have short A tails. Nucleic Acids Res. **27:**2416–2422.
13. **Gutierrez, R. A., G. C. MacIntosh, and P. J. Green.** 1999. Current perspectives on mRNA stability in plants: multiple levels and mechanisms of control. Trends Plant Sci. **4:**429–438.

14. **Hanic-Joyce, P. J., and M. W. Gray.** 1990. Processing of transfer RNA precursors in a wheat mitochondrial extract. J. Biol. Chem. **265:**13782–13791.
15. **Hayes, R., J. Kudla, and W. Gruissem.** 1999. Degrading chloroplast mRNA: the role of polyadenylation. Trends Biochem. Sci. **24:**199–202.
16. **Kaleikau, E. K., C. P. Andre, and V. Walbot.** 1992. Structure and expression of rice mitochondrial apocytochrome b gene (*cob*-1) and pseudo-gene (*cob*-2). Curr. Genet. **22:**463–470.
17. **Kudla, J., R. Hayes, and W. Gruissem.** 1996. Polyadenylation accelerates degradation of chloroplast mRNA. EMBO J. **15:**7137–7146.
18. **Lisitsky, I., P. Klaff, and G. Schuster.** 1996. Addition of destabilizing poly(A)-rich sequences to endonuclease cleavage sites during the degradation of chloroplast mRNA. Proc. Natl. Acad. Sci. USA **93:**13398–13403.
19. **Lupold, D. S., A. G. F. S. Caoile, and D. B. Stern.** 1999. Polyadenylation occurs at multiple sites in maize mitochondrial *cox2* mRNA and is independent of editing status. Plant Cell **11:**1565–1577.
20. **Marchfelder, A., and A. Brennicke.** 1994. Characterization and partial purification of tRNA processing activities from potato mitochondria. Plant Physiol. (Rockville) **105:**1247–1254.
21. **Morikami, A., and K. Nakamura.** 1993. Transcript map of oppositely oriented pea mitochondrial genes encoding the α-subunit and the subunit 9 of F_0F_1-ATPase complex. Biosci. Biotechnol. Biochem. **57:**1530–1535.
22. **Ojala, D., J. Montoya, and G. Attardi.** 1981. tRNA punctuation model of RNA processing in human mitochondria. Nature **290:**470–474.
23. **Saalaoui, E., S. Litvak, and A. Araya.** 1990. The apocytochrome b from an alloplasmic line of wheat (*T. aestivum*, cytoplasm-*T. timopheevi*) exists in two differently expressed forms. Plant Sci. **66:**237–246.
24. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
25. **Schuster, D., I. Lisitsky, and P. Klaff.** 1999. Polyadenylation and degradation of mRNA in the chloroplast. Plant Physiol. (Rockville) **120:**937–944.
26. **Schuster, W., R. Hiesel, P. G. Isaac, C. J. Leaver, and A. Brennicke.** 1986. Transcript termini in messenger RNAs in higher plant mitochondria. Nucleic Acids Res. **14:**5943–5954.
27. **Wahle, E., and U. Rüegsegger.** 1999. 3′-end mRNA processing of pre-mRNA in eukaryotes. FEMS Microbiol. Rev. **23:**277–295.
28. **Yehudai-Resheff, S., and G. Schuster.** 2000. Characterization of the *E. coli* poly(A) polymerase: nucleotide specificity, RNA binding affinities and RNA structure dependence. Nucleic Acids Res. **28:**1139–1144.
29. **Yuckenberg, P. D., and S. L. Phillips.** 1982. Oligoadenylate is present in the mitochondrial RNA of *Saccharomyces cerevisiae*. Mol. Cell. Biol. **2:**450–456.

6.1.3 5’ meets 3’: head to tail linked mRNAs in plant mitochondria

5’ meets 3’: head to tail linked mRNAs in plant mitochondria

Josef Kuhn and Stefan Binder*

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

ABSTRACT

A recent analysis of the 3’ ends of the pea *atp9* mitochondrial transcripts suggested the occurrence of mRNAs with 5' to 3' linked termini. Detailed RT-PCR analysis of pea *atp9* and *Oenothera atp1* mRNAs now reveals head to tail connected transcripts in the steady state mRNAs in plant mitochondria. In the cDNA clones deriving from *in vivo Oenothera atp1* mRNA, one 3’ end was detected exclusively in the head to tail linked transcripts. The other borders of the linked RNA moieties coincide with 5' and 3' termini of transcripts colinear with genomic sequences. The analysis of head to tail linked transcripts reveals a novel 5’ end of *Oenothera atp1* RNAs. This is located within a 21 nucleotides long sequence segment also conserved at the mature 5’ ends of 18S rRNAs of several plant species. Head to tail transcripts in pea and *Oenothera* mitochondria enclose short stretches of up to five non-encoded adenosines. The frequency of non-encoded nucleotides in these cDNA clones suggests that head to tail connection of transcripts is independent from their addition. The observed covalent linkages between the 5' and 3' termini suggest the presence of a ligase activity towards RNA in plant mitochondria.

INTRODUCTION

The mechanisms controlling the expression of the genetic information of plant mitochondria are important components of the total cellular regulatory network. In recent years it has become apparent that beside the level of transcription initiation posttranscriptional processes play a substantial role in the regulation of mitochondrial genes (1, 2). Particularly the processes influencing RNA stability are major determinants of the steady state level of certain if not all transcripts. An important role in RNA stability has been attributed to 3’ terminal stem-loop structures (3, 4). For instance a double stem-loop structure at the 3' end of the pea *atp9* transcripts does not terminate transcription *in vitro* but more likely plays a significant role in the posttranscriptional formation of the mature 3' terminus of this RNA (4). Its protective impact against degradation is high enough to largely neutralize the degradation promoting effect of poly(A) tails (5). Polyadenylation has been found at several mitochondrial mRNAs from different plant species and *in vitro* data suggest that poly(A) extensions, whose sizes *in vivo* are still a mystery, generally accelerate degradation by targeting the RNA for specific degradation by most likely exonucleolytic ribonuclease activities (6-8). Such destabilizing effects of poly(A) tails have also been reported in *E. coli* and chloroplasts, but here substantially stronger enhancing the decay of the polyadenylated transcripts (9-11). Conversely completely different functions are attributed to polyadenylation of mRNAs in the nuclear/cytoplasmic compartments of eukaryotic cells. Here poly(A) tails support the stability of mRNAs and play an important role in the export of mRNA from the nucleus to the cytoplasm (12). Furthermore the close interaction of the poly(A) sequences with the 5’ cap results in a synergistic enhancement of translation (13). In yeast this important communication is mediated by several RNA-protein and protein-protein interactions involving the cap binding protein eIF4E, factor eIF4G and the poly(A) binding protein Pab1b forming a pseudo-circular form of the mRNA. Despite the clearly translation stimulating effect the exact function of this structure is unclear at present, but is interpreted to increase the stability of the RNA or to have some advantage in ribosome recycling for multiple translation of an mRNA molecule (13).

In our studies of posttranscriptional control of gene expression in plant mitochondria we focus on the formation of mature 3' transcript termini. In an analysis of the 3' end of pea *atp9* transcripts we detected two clones deriving from RNA molecules, in which the 5’ end of the shortest of the three mRNAs is covalently linked to the mature 3' end (5). Although RNA ligase was used in this experimental approach, the origin of these cDNA clones from independent experiments and even

*Corresponding author:
Stefan Binder
Molekulare Botanik, Universität Ulm,
Albert-Einstein-Allee 11, 89069 Ulm,
Germany
Phone: +49 731 502 2625
Fax: +49 731 502 2626
E-mail: stefan.binder@biologie.uni-ulm.de

different mitochondrial RNA preparations let us speculate that such 5' to 3' end connected pea *atp9* transcripts may exist *in vivo*. We now report that *atp9* transcripts with head to tail conformation are indeed present *in vivo* in pea mitochondria. Analogously 5' to 3' linked molecules are also found to be derived from the *Oenothera atp1* gene. This is to our knowledge the first report of head to tail linked transcripts in mitochondria.

MATERIALS AND METHODS

RT-PCR analysis of mitochondrial RNA

Total mitochondrial RNA was isolated from six to seven days old etiolated pea seedlings and from *Oenothera* callus culture as described previously (14). For the detection of *in vivo* head to tail joined RNA molecules first strand synthesis was carried out on 9 µg total mitochondrial RNA (mtRNA) initiated with oligonucleotides PA3 (335 to 319, pea *atp9*) and OA3 (376 to 356 *Oenothera atp1*), respectively. The reactions were performed with Superscript reverse transcriptase (RT) under conditions recommended by the manufacturer (InVitrogen Life Technologies). The cDNAs were then used for the amplification by PCR with Klentaq polymerase following the manufacturer's manual (Clontech). Pea *atp9* cDNAs were amplified with primers PA2 (291 to 274) and T7IR+ (380 to 395) in six cycles of 1 min at 94°C, 1 min at 50°C and 30 sec at 68°C followed by 35 cycles with 1 min at 94°C, 1 min at 44°C and 30 sec at 68°C. Nested PCRs were performed under identical conditions with primer pairs PA19 (22 to 5) / PAX1 (396 to 411) and PA21 (260 to 242) / PAX1. Oligonucleotide pairs OA2 (302 to 285) / ORace1 (1571 to 1589) and OA4 (252 to 232) / OAX1 (1930 to 1945) were used to generate *Oenothera atp1* cDNAs following the same experimental regime. All PCR products were cloned into an *Xcm*I-linearized pBluescript II vector and sequenced as detailed below. The location of the oligonucleotides is given in respect to the first nucleotide of the translation start codon (+).

Miscellaneous methods

Standard methods in molecular biology were performed as described elsewhere (15). Primer extension experiments were performed with primers OA2 and 18S-2 (104 to 87 relative to the mature 5' end) as explained in previous analyses (16). Thermo Sequenase fluorescent labeling kits and Cy5 labeled primers were used for DNA sequence analysis and sequencing products were separated and analyzed with ALF express sequencers (Amersham Biosciences). *In vitro* ligation reactions with T4 RNA ligase were performed as described previously (17).

RESULTS

Detection of 5' to 3' end ligated *atp9* mRNAs in pea mitochondria

By ligating an anchor primer directly to 3' ends we recently investigated the *in vivo* 3' termini of pea *atp9* transcripts (5). The anchor primer serves as a binding site for a complementary oligonucleotide from which cDNA synthesis was initiated. This oligonucleotide in combination with gene-specific primers was used in subsequent amplification reactions. The resulting PCR fragments, which enclose the 3' terminal sequences, were then inspected by cloning and DNA sequencing. In this approach we detected two clones, which derived from transcripts where 5' moieties of pea *atp9* mRNA are connected to the 3' parts of the same transcript, the linker being ligated at the 3' end of the 5' derived fragment. The mRNA linkage is exactly between the mature 3' and 5' termini of the shortest *atp9* mRNA as determined by primer extension analysis (18). Since these clones were obtained in independent ligation and amplification reactions from different RNA preparations, it seemed unlikely that 5' and 3' parts of the same transcript are incidentally connected *in vitro* by the T4 RNA ligase used for the covalent addition of the anchor oligonucleotide to the 3' termini of the RNA.

To investigate whether head to tail ligated pea *atp9* transcripts exist *in vivo*, an RT-PCR analysis was carried out, in which the cDNA synthesis is initiated from sequences of the 5' region of the mRNA followed by an amplification reaction across the potentially 5' to 3' connected ends (Fig. 1, A). The cDNA synthesis was initiated from primer PA3 on total mitochondrial RNA. The first amplification reaction generated major products of about 250 and 350 bp, several minor fragments and a background smear ranging up to 1.0 kb (Fig. 1B, left panel). Estimated from the previously mapped 5' and 3' ends a product of about 390 bp is expected upon linkage of intact termini. Thus cDNA fragments with sizes between 300 and 500 bp were recovered from the gel and used as templates in two further amplification reactions with alternative nested primer pairs. A PCR with oligonucleotides PAX1 and PA19 generated a major product of about 80 bp consistent with the expected size (Fig. 1B, middle panel, lane 2a). In the PCR with primers PAX1 and PA21 a relatively weak fragment with the anticipated size of about 330 bp is observed (Fig. 1B, right panel, lane 2b). The amplification of products with sizes expected indicated the presence of head to tail connected pea *atp9* RNAs and thus both fragments were cloned and sequenced.

***In vivo* 5' to 3' end linkage is independent from the presence of non-encoded nucleotides at the junctions**

Sequence analysis of 29 clones confirmed the *in vivo* existence of 5' to 3' end linked pea *atp9* mRNAs. In 15 clones the observed termini of the 5' moiety coincide with the 5' terminal nucleotide (-136 relative to the ATG) of the shortest pea *atp9* mRNA (18)(Fig. 2A and B). In five clones the regions corresponding to the 5' part begin at position -135 and the 5' moieties of the other clones start between -134 and –12. These are, with the exception of two clones starting with position –121 found only in single cDNAs. In 22 clones the 5' parts are joined to 3' ends located directly downstream of a double stem-loop or within the downstream part of the second stem. The other cDNAs revealed 3' ends within the downstream loop and a single clone contains a 3' terminus in the first stem. The scattering of the 3' ends has similarly been observed in previous experiments with an analogous distribution of the termini (4, 5, 18).
The sequences not only reveal *in vivo* head to tail connected pea *atp9* transcripts, they also display the presence of non-encoded nucleotides in these RNAs. In 16 clones one to four nucleotides are embedded between the 5' and 3' moieties. The majority of these additions are primarily two or three adenosines, while a cytidine and a guanosine are found in single clones. Such short sequences have previously been detected at the 3' termini of "normal" linear pea *atp9* transcripts (5). Since there is also a substantial amount of clones without any non-encoded nucleotides, there is no apparent correlation between the 5' to 3' joining and the presence of non-encoded nucleotides. In all clones derived from the PA21/PAX1 PCR fragments, cytidine to uridine exchanges are found at positions 20 and 50 relative to the first nucleotide of the translation start codon (+1). These editing sites have also been identified in cDNAs originating from linear RNAs (data not shown).

Head to tail *atp1* transcripts are present in Oenothera mitochondria

To follow the *in vivo* presence of head to tail connected transcripts on a wide basis, we extended our analysis to *atp1* transcripts from *Oenothera* mitochondria. To this end a similar RT-PCR strategy was used in order to amplify cDNA fragments from *atp1* RNAs with an analogous structure. Following first strand cDNA synthesis from oligonucleotide OA3 major cDNA fragments with sizes between 70 and 150 bp were amplified with oligonucleotides OA2 and ORaceI (Fig. 3B, left panel, lane 1). Minor products and a smear ranging up to more than 750 nucleotides were also observed. Since a product of 720 nucleotides is calculated from previously determined 5' and 3' ends (14, 19), nucleic acids with sizes between 500 and 800 nucleotides were used as templates in a second PCR with nested oligonucleotides OA4 and OAXI. This generated two prominent products of about 200 and 250 nucleotides, respectively, both slightly shorter than the 320 nucleotides expected (Fig. 3B, right panel, lane 2). Similar results were obtained in a control reaction with total mitochondrial RNA that had been self ligated *in vitro* with T4 RNA ligase. Here a PCR fragment of the expected size was already detectable after the first PCR and a single major product of 250 nucleotides was obtained in the second PCR (data not shown). From both approaches cDNA fragments between 180 and 270 nucleotides were cloned and sequenced.
From 32 clones obtained from the approach without T4 RNA ligase only a single cDNA contained a 5' moiety starting at the transcription initiation site (Fig. 4A and B, position -209). The majority of the clones (27) contain 5' parts beginning with positions -131 through –135. This suggests new 5' ends, which map downstream of the transcription initiation sites and which have not been identified in previous analyses (14, 19). The termini of the 5' parts of the remaining four clones coincide with nucleotides -199, -72, -9 and -8.
The ends of the 3' moieties detected in this approach can be divided into two categories. The first collection scatters over a few nucleotides directly downstream of the double stem-loop with the 3' end of a single clone located at the beginning of the downstream part of the first stem. These ends found in 22 clones are consistent with previously mapped 3' ends and also coincide with poly(A) addition sites detected in a 3' RACE analysis initiated with an oligo(dT)-adapter primer (5, 20). The second category of 3' termini maps within the upstream loop of the double stem-loop. These ends are represented by ten cDNA clones and have not been described before. This division of the 3' ends is further suggested by the presence or absence of non-encoded nucleotides. While four and five, respectively, non-encoded adenosines are found in 16 of the cDNAs with 3' moieties ending downstream of the double stem-loop structure, no unambiguously non-encoded nucleotides are seen in clones with the shorter 3' termini (Fig. 4B).
A more homogenous pattern of 5' and 3' borders is found in clones amplified from *in vitro* ligated mitochondrial RNA (Fig. 4C). Here all but one clones match the –131 to -136 5' ends. Likewise almost all 3' sequences end directly downstream of the double stem-loop. Only a single cDNA contains a 3' end located within the downstream part of the second stem. In contrast the length of the non-encoded adenosine stretches found in this collection of clones is slightly more variable, ranging from one to five nucleotides (Fig. 4C).
Since 5' termini at -131 to -136 have not been detected previously these ends were further investigated by an independent experimental

approach. To this end a primer extension analysis with oligonucleotide OA2 annealing downstream of the translation start codon was performed. Two strong extension products of about 230 and 300 nucleotides, respectively, indeed correlate with the here detected 5' ends at position -131 through -136 and the previously mapped transcription initiation site, respectively (position -209, Fig. 4A and 5). Thus the –131 to -136 ends are also major 5' termini of transcripts with a "normal" linear conformation.
In summary the detection of head to tail connected conformation is not restricted to a single gene in a single plant species but seems to be a frequent phenomenon in plant mitochondria.

DISCUSSION

Head to tail linked transcripts represent a subpopulation of plant mitochondrial steady state RNA

The analysis of pea *atp9* and *Oenothera atp1* mitochondrial transcripts by RT-PCR analysis reveals the presence of head to tail connected RNA molecules in plant mitochondria. The fact that the ends of the 5' moieties most often coincide with major 5' ends of RNAs colinear to DNA makes an RT-PCR artifact as the origin for the here detected cDNA fragment highly unlikely (14, 19, 20). In addition, transcripts with head to tail conformations are not detectable for every gene since no such RNAs could be obtained from the pea *cox2* gene. For these RNAs head to tail conformations are only detectable after *in vitro* ligation with T4 RNA ligase demonstrating that the PCR strategy applied should be compatible with the detection of potential *in vivo* head to tail connected RNAs (17). At the same time this result confirms the validity of the strategy to detect genuine in vivo 5' - 3' conformation and excludes an artificial generation of such molecules during the experimental procedures.
The generation of cDNA fragments from *in vivo* head to tail connected transcripts only after two rounds of PCR with nested primers indicates that they represent a minor fraction in the mitochondrial steady state RNA. A preliminary RT-PCR quantification of pea *atp9* of RNAs with head to tail and "normal" linear conformations suggests that the former represent less than 1 % of the total *atp9* mRNA pool (data not shown).

Linear versus circular conformation of the head to tail connected mitochondrial mRNAs

The here applied PCR strategy is designed to detect head to tail connected transcripts whose presence was suspected from the previous 3' RACE analysis. This strategy, however, will amplify both circular and linear head to tail connected transcripts and does not discriminate between the two forms as the origin for the cDNA fragments. The existence of linear head to tail forms of pea *atp9* mRNAs has been demonstrated in previous 3' RACE experiments including *in vitro* ligation of an anchor-oligonucleotide to free hydroxyl groups at 3' transcript termini, a reaction only possible with non-circular RNAs (5). The two clones detected contain 24 nucleotides long 5' terminal parts connected to the 3' end. In the now applied experimental strategy the cDNA initiated at oligonucleotide PA3 covers 336 of about 445 nucleotides of the complete *atp9* RNA demonstrating that up to 75% of the complete mRNA can be connected via the 5' – 3' junction. Linkage of such large parts of the RNA indicates that circular transcripts exist *in vivo,* although there is no direct experimental evidence for this.
Circular RNAs in mitochondria have been found in 11 S and 18 S RNAs in yeast mitochondria (21). These are transcripts of the *oxi3* and *oxi1* gene loci and were later identified as closed circular forms of excised introns (22). Such molecules actually lariats formed in the splicing process of group II introns, are also present in mitochondria of higher plants. These lariats are formed by two transesterification reactions and the formation of a 2'-5' bond between the 5' end of the intron and a so-called bulged adenosine located a few nucleotides upstream of the 3' end (23, 24). However, to our knowledge, no circular form of any RNA was so far detected from mitochondrial genes without introns as the here investigated *atp9* and *atp1* genes.
Unclear remains at present the mechanism leading to the formation of head to tail connected conformations. We assume that the observed connection is generated via a circularization since an intramolecular covalent linkage of 5' and 3' ends is more probable than an intermolecular connection. Circular forms could subsequently be linearized by an endonucleolytic cut or mechanical shearing forces generating linear molecules with head to tail constitutions.

Correlation of 3' ends and non-encoded nucleotides between head to tail and linear transcripts

In *Oenothera atp1* head to tail connected transcripts the 3' ends located in the upstream inverted repeat are detected in about one third of the inspected clones. Analogous ends are not found in clones originating from *in vitro* ligated RNA or in "normal" linear RNAs and thus seem to be restricted to the *in vivo* head to tail connected molecule conformation. All other ends essentially coincide with the termini found at RNAs with the "normal" linear structure (5, 19).
Similarities between the two conformations exist also in terms of the frequency, lengths and

composition of non-encoded nucleotides. Primarily short extensions of four or five adenosines are found in both conformations of the *Oenothera atp1* transcripts. Slightly shorter predominantly homoadenosine extensions are present in the pea *atp9* head to tail conformations. Such lengths are consistent with the extensions detected in the 3' RACE analysis. In the latter, however, other nucleotide identities, predominantly cytidines are more frequently observed. Similar short extensions have also been found at several maize mitochondrial transcripts (25). The function of the short extensions detected in 30 to 50 % of the clones is still unclear. *In vitro* assays showed that such short tails of three or five adenosines do not accelerate degradation (5, 6). They most likely represent relatively stable intermediates generated during synthesis of larger tails or are byproducts of a stalling or interrupted degradation. The latter explanation is substantiated by previous *in vitro* analyses, in which the double stem-loop interferes with the complete degradation of a transcript and polyadenylation promotes only a rapid removal of downstream nucleotides (5).
A large proportion of the RNAs with head to tail conformation does not contain any non-encoded nucleotides suggesting that the formation of this molecule structure does not require nor exclude the presence of non-encoded nucleotides.

The sequence surrounding the newly identified *Oenothera atp1* 5' end is similar to the processing site at the mature 5' termini of 18S rRNAs

The analysis of *in vivo* and *in vitro* head to tail connected *Oenothera atp1* transcripts revealed new 5' ends about 75 nucleotides downstream of the transcription initiation site. A 21 nucleotide sequence segment comprising these termini is also found at 5' ends of the mature 18S rRNA in different plant species (Fig. 6) (26). A primer extension analysis of this rRNA from *Oenothera* further showed these ends to be at similar nucleotide positions as those of the *atp1* mRNAs. Since the mature 18S rRNA 5' end is generated by processing (most likely endonucleolytic), a similar scenario is likely for the newly identified –131 to –136 *atp1* ends (Fig. 6, left panel). The short identical sequence conserved at these ends could serve as a processing signal for an endonucleolytic cleavage. A different conserved block of 24 nucleotides is found at the mature 5' end of *Oenothera* 26S rRNA and the 5' terminus of the shortest transcript of the *atp9* gene in the same organism (27). This may suggest that endonucleases generating the mature termini of the ribosomal RNAs also operate on mRNAs containing identical or very similar sequence segments. The primary sequences rather than higher order structures of the RNAs may thus determine the site of the endonucleolytic cleavage. The function and significance of the processing within the 5' untranslated region of the *atp1* mRNAs remains unclear, since only about half of the steady state *atp1* transcripts are processed as estimated from the primer extension analysis (Fig. 5).

Is an RNA ligase involved in the formation of head to tail connected RNA species?

One of the most important questions is how the extremities of an RNA are connected *in vivo*. The most likely explanation is the joining by an RNA ligase and there are indeed arguments that point into this direction.
First, only a single cDNA clone deriving from *in vivo* connected *Oenothera atp1* transcripts exactly coincides with the transcription initiation site, while 27 of 32 clones (84%) contain the –131 to –136 ends. A similar bias is observed among the 5' ends detected after *in vitro* ligation with T4 RNA ligase. Here 20 of 21 clones contain the –131 to –136 termini and only a single clone corresponds to an RNA with a 5' end close to the transcription initiation site. These frequencies are in contrast to ends found in the primer extension analysis. Here similar amounts of the steady state *atp1* transcripts have 5' ends corresponding to the transcription start site and to the –131 to –136 ends, respectively (Fig. 5). Thus head to tail connection *in vivo* and *in vitro* by T4 RNA ligase have a clear bias towards the shorter end. This preference for the shorter ends is determined by the ligase reaction mechanism, which allows only the ligation of 5' ends with monophoshpate groups suggesting that 5' to 3' end joining *in vivo* is also performed by an RNA ligase.
Second, in pea mitochondria no *in vivo* head to tail connected *cox2* transcripts are detectable. For this gene three different 5' termini have been mapped and there is strong evidence that all of these originate from transcription initiation events and thus carry 5' triphosphates. Consistent with this 5' ends in cDNAs from *in vitro* ligated RNAs never coincide with the primary 5' termini mapped by primer extension analysis (17). Only a subpopulation of RNAs with slightly shortened 5' ends are ligated. Thus the amount of *in vivo* head to tail RNA species that could be formed by an RNA ligase may be extremely low escaping detection by RT-PCR.
Mitochondrial RNA ligases have been identified in kinetoplastida, where they play a crucial role in the posttranscriptional RNA editing process (28, 29). In these protists two such enzymes are involved in the ligation of the mRNA moieties after insertion or deletion of uridine(s). They are also responsible for the ligation of the mRNA upstream part to guide RNAs (30), but there is to our knowledge no report of head to tail ligated mRNAs in kinetoplastida. Ligases operating on DNA are present in mitochondria of human, mouse, *Xenopus* and yeast (31-34). These enzymes are suggested to be

involved in replication, DNA repair and possibly also in recombination. Nothing is known about RNA or DNA ligases in plant mitochondria. While DNA replication definitively requires the presence of a DNA ligase, the here reported linkage of mRNA termini is the first report on a process requiring RNA ligase in plant mitochondria.

What is the function of head to tail connected transcripts in plant mitochondria?

At present one can only speculate about the importance of transcripts with a head to tail conformation. They are clearly a minor fraction of the steady state RNA pool indicating that a potentially increased stability through protection against exonucleolytic degradation might not be an important feature. Such protection would only be yielded by the circular form, which so far has not been observed *in vivo*. The circular form could resemble the pseudo circular form of nuclear/cytoplasmic mRNAs with a non-covalent interaction between 5' and 3' terminus mediated by several proteins enhancing translation efficiency (13, 35). Although it remains to be determined whether a similar effect can be attributed to covalently connected circular mitochondrial transcripts, the here detected head to tail conformation may indicate an interaction of the 5' and 3' extremities of at least certain mRNAs even if the formation of the phosphodiester bond occurs incidentally in plant mitochondria.

ACKOWLEDGEMENTS

We very much appreciate the excellent technical assistance of Bärbel Weber. This work was supported by grants Bi 590/3-3 and 3-4 to S.B. and Fonds der Chemischen Industrie. J.K. is a fellow of the Studienstiftung des Deutschen Volkes.

REFERENCES

1. Giege, P., Hoffmann, M., Binder, S. & Brennicke, A. (2000) *EMBO Rep.* **1,** 164-170.
2. Hoffmann, M., Kuhn, J., Daschner, K. & Binder, S. (2001) *Prog. Nucleic Acid Res. Mol. Biol.* **70,** 119-154.
3. Bellaoui, M., Pelletier, G. & Budar, F. (1997) *EMBO J.* **16,** 5057-5068.
4. Dombrowski, S., Brennicke, A. & Binder, S. (1997) *EMBO J.* **16,** 5069-5076.
5. Kuhn, J., Tengler, U. & Binder, S. (2001) *Mol. Cell. Biol.* **21,** 731-742.
6. Gagliardi, D., Perrin, R., Marechal-Drouard, L., Grienenberger, J. M. & Leaver, C. J. (2001) *J. Biol. Chem.* **276,** 43541-43547.
7. Gagliardi, D. & Leaver, C. J. (1999) *EMBO J.* **18,** 3757-3766.
8. Lupold, D. S., Caoile, A. G. & Stern, D. B. (1999) *Plant Cell* **11,** 1565-1578.
9. Carpousis, A. J., Vanzo, N. F. & Raynal, L. C. (1999) *Trends Genet.* **15,** 24-28.
10. Hayes, R., Kudla, J. & Gruissem, W. (1999) *Trends Biochem. Sci.* **24,** 199-202.
11. Schuster, G., Lisitsky, I. & Klaff, P. (1999) *Plant Physiol.* **120,** 937-944.
12. Wahle, E. & Ruegsegger, U. (1999) *FEMS Microbiol. Rev.* **23,** 277-295.
13. Wells, S. E., Hillner, P. E., Vale, R. D. & Sachs, A. B. (1998) *Mol. Cell* **2,** 135-140.
14. Binder, S. & Brennicke, A. (1993) *J. Biol. Chem.* **268,** 7849-7855.
15. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).
16. Marchfelder, A., Brennicke, A. & Binder, S. (1996) *J. Biol. Chem.* **271,** 1898-1903.
17. Kuhn, J. & Binder, S. (2002) *Nucleic Acids Research* **30,** 439-446.
18. Morikami, A. & Nakamura, K. (1993) *Biosci. Biotechnol. Biochem.* **57,** 1530-1535.
19. Schuster, W. (1988) PhD thesis *Fakultät für Biologie* (Eberhard-Karls-Universität Tübingen, Tübingen).
20. Schuster, W., Hiesel, R., Isaac, P. G., Leaver, C. J. & Brennicke, A. (1986) *Nucleic Acids Res.* **14,** 5943-5954.
21. Arnberg, A. C., Van Ommen, G. J., Grivell, L. A., Van Bruggen, E. F. & Borst, P. (1980) *Cell* **19,** 313-319.
22. van der Veen, R., Arnberg, A. C., van der Horst, G., Bonen, L., Tabak, H. F. & Grivell, L. A. (1986) *Cell* **44,** 225-234.
23. Qin, P. Z. & Pyle, A. M. (1998) *Curr. Opin. Struct. Biol.* **8,** 301-308.
24. Michel, F. & Ferat, J. L. (1995) *Annu. Rev. Biochem.* **64,** 435-461.
25. Williams, M. A., Johzuka, Y. & Mulligan, R. M. (2000) *Nucleic Acids Res.* **28,** 4444-4451.
26. Giese, A., Thalheim, C., Brennicke, A. & Binder, S. (1996) *Mol. Gen. Genet.* **252,** 429-436.
27. Schuster, W. & Brennicke, A. (1989) *Curr. Genet.* **15,** 187-192.
28. Schnaufer, A., Panigrahi, A. K., Panicucci, B., Igo, R. P., Jr., Wirtz, E., Salavati, R. & Stuart, K. (2001) *Science* **291,** 2159-2162.
29. Rusche, L. N., Huang, C. E., Piller, K. J., Hemann, M., Wirtz, E. & Sollner-Webb, B. (2001) *Mol. Cell. Biol.* **21,** 979-989.
30. Sabatini, R. & Hajduk, S. L. (1995) *J. Biol. Chem.* **270,** 7233-7240.
31. Lakshmipathy, U. & Campbell, C. (1999) *Mol. Cell. Biol.* **19,** 3869-3876.
32. Perez-Jannotti, R. M., Klein, S. M. & Bogenhagen, D. F. (2001) *J. Biol. Chem.*, 48978-48987.
33. Mackey, Z. B., Ramos, W., Levin, D. S., Walter, C. A., McCarrey, J. R. & Tomkinson, A. E. (1997) *Mol. Cell. Biol.* **17,** 989-998.
34. Willer, M., Rainey, M., Pullen, T. & Stirling, C. J. (1999) *Curr. Biol.* **9,** 1085-1094.
35. Sachs, A. B. & Varani, G. (2000) *Nat. Struct. Biol.* **7,** 356-361.

FIGURE LEGENDS

Fig. 1. Detection of head to tail joined *atp9* transcripts in pea mitochondria. A, Strategy used for the detection of *in vivo* 5' end to 3' end ligated pea *atp9* mRNAs. Oligonucleotides used for the RT-PCR analysis across the *in vivo* junction are indicated by horizontal arrows along the mRNA (bold line). The *atp9* reading frame is bordered by the ATG and TAA codons. The orientation of the transcript is indicated by fading from black (5' end) to white (3' end). The potential double stem-loop structure at the 3' terminus is indicated. For clearer presentation of the location and orientation of the oligonucleotides, the mRNA is shown in circular form, although there is no experimental evidence for this conformation. B, Agarose gel electrophoresis of PCR products obtained in the RT-PCR analysis across the 5' -3' junction. Two major products of about 350 and 200 nucleotides are obtained in the first reaction with primers PA2 and T7IVR+ (left panel). Since products of about 390 nucleotides are expected upon the joining of the major 5' and 3' termini, cDNAs with sizes between 250 and 500 bp (range indicated on the right) are recovered from the gel and used as DNA templates in two nested PCRs with oligonucleotide pairs PA19/PAXI (center panel) and PA21/PAXI (right panel). Fragments of about 80 and 330 bp, respectively, are cloned and sequenced (lanes 2a and 2b). Sizes of coelectrophoresed DNA marker fragments (lanes M) are given on the left hand sides.

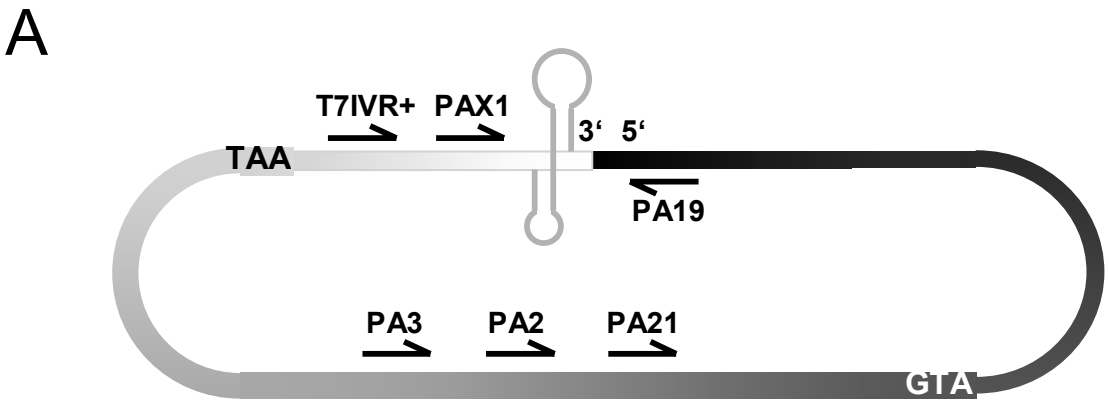
Fig. 2. cDNA sequences across pea *atp9 in vivo* 5' to 3' junctions. A, Schematic structure of the *atp9* locus in the pea mitochondrial genome. Three 5’ ends (bent arrows) and a single 3’ terminus (inverted arrow heads) have been mapped upstream and downstream of the pea *atp9* reading frame (box). The genomic sequences surrounding the 5’ end of the shortest transcript and the 3’ end are indicated. Vertical arrows beneath the sequence indicate transcript 3’ termini found in previous analyses. B, Sequences of 29 individual cDNA clones. The number of clones found for each of the 5’ to 3’ linkages (dash) is given on the left. 5' sequences are given in white letters on black fading boxes. Non-encoded nucleotides are given in lower case letters. Numbering refers to the first nucleotides of the translation start codon (+1). Parts of the double inverted repeat are indicated by gray arrows in the background of the 3’ sequence.

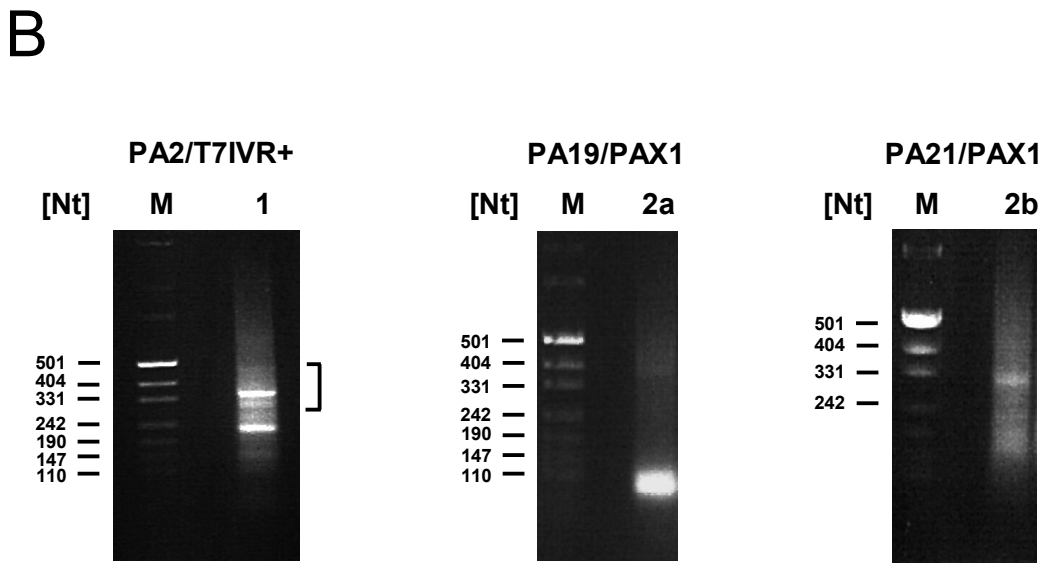
Fig. 3. Analysis of *Oenothera atp1* transcripts. A, Schematic representation of the *atp1* mRNA in the circular form and location of the primers used. B, Agarose gel analysis of PCR products obtained with primer pairs OA2/ORacel (left panel) and OA4/OAX1 (right panel). Since a product of about 730 nucleotides is expected in the first PCR (lane 1) cDNAs with sizes between 500 through 800 nucleotides (range indicated by bracket) are used as DNA templates in the second PCR. This generated a fragment of about 250 nucleotides (lane 2), about 70 nucleotides shorter than expected from the single previously mapped 5’ terminus.

Fig. 4. *Oenothera atp1* cDNA sequences of head to tail connected 5' and 3' extremities. A, Genomic sequences at 5’ and 3’ termini of *Oenothera atp1* transcripts. B and C. Sequences encompassing of head to tail connection of *in vivo* (B) and *in vitro* (C) joined RNAs. Structure, marking and lettering is analogous to Fig. 2B.

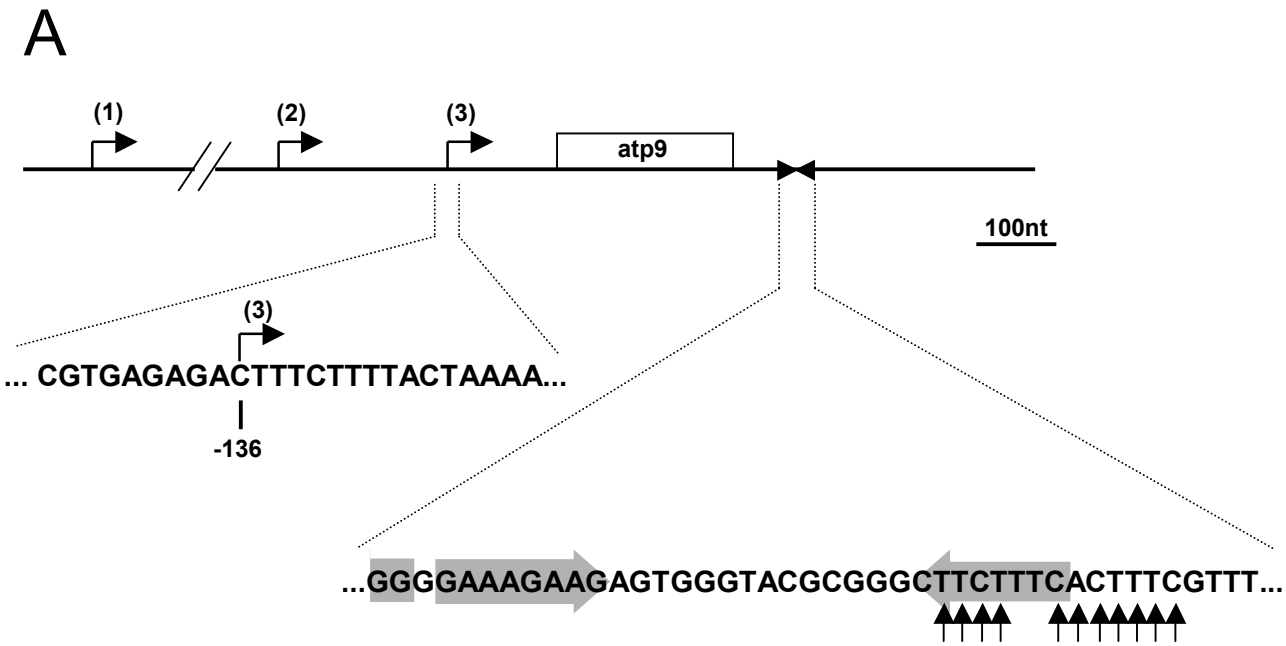
Fig. 5. Primer extension analysis of *Oenothera atp1* transcripts. The upper part shows primer extension products (lane P, numbered arrows) separated alongside pGEM marker fragments (lane M), the sizes of which are given in nucleotides at the left. The numbering of the extension products corresponds to the cartoon. Bottom part: First strand synthesis is initiated at oligonucleotide OA2 (location indicated by a vertical line) annealed within the *atp1* coding region (open box). The larger of two extension products (dashed lines) marks the transcription initiation site indicated by the bent arrow on the hatched box.

Fig. 6. Sequence box conserved at *Oenothera atp1* and 18S rRNA 5’ ends. Left panel: Primer extension products (lane P) obtained with oligonucleotide 18S-2 were coelectrophoresed with sequencing fragments (G, A, T, C) obtained with the same primer. Left panel: The –131 through –136 5' ends (horizontal arrows, lengths indicating the relative abundance) of *Oenothera atp1* transcripts are found at similar nucleotide positions within a sequence element (boxed) at the mature 5' termini of the 18S rRNA genes (*rrn*18) from *Oenothera berteriana* (O. b.), *Solanum tuberosum* (S. t.), *Pisum sativum* (P. s.) and *Arabidopsis thaliana* (A. t.).





Kuhn and Binder, Fig. 1



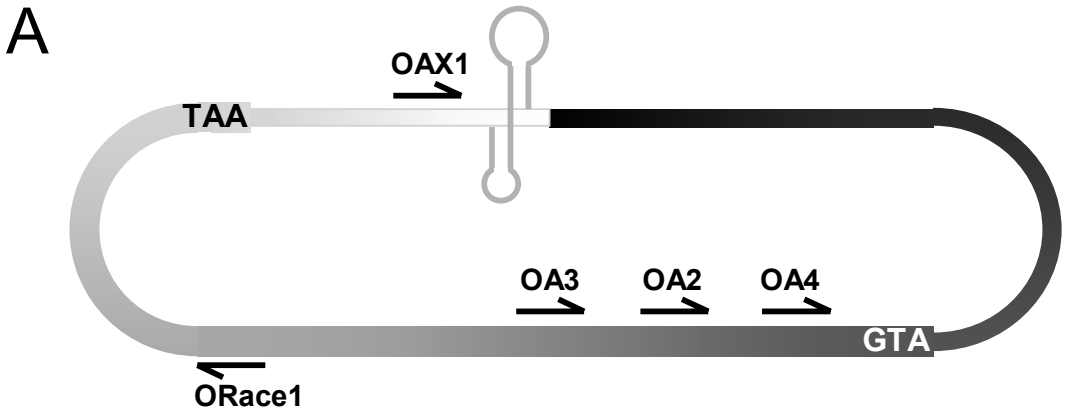
Kuhn and Binder, Fig. 2

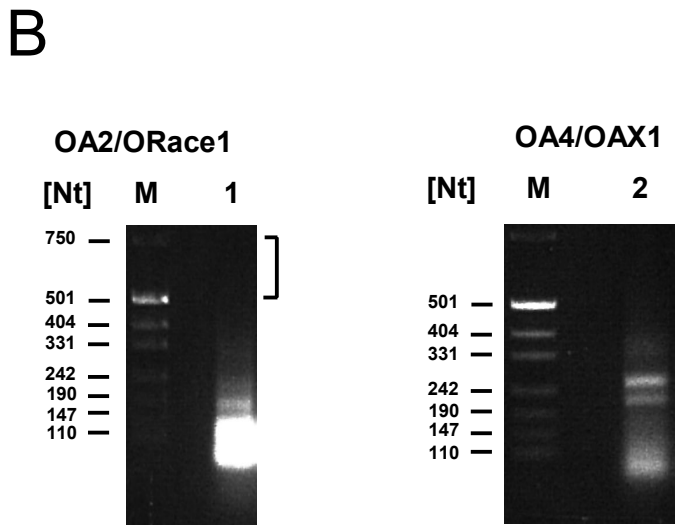
B



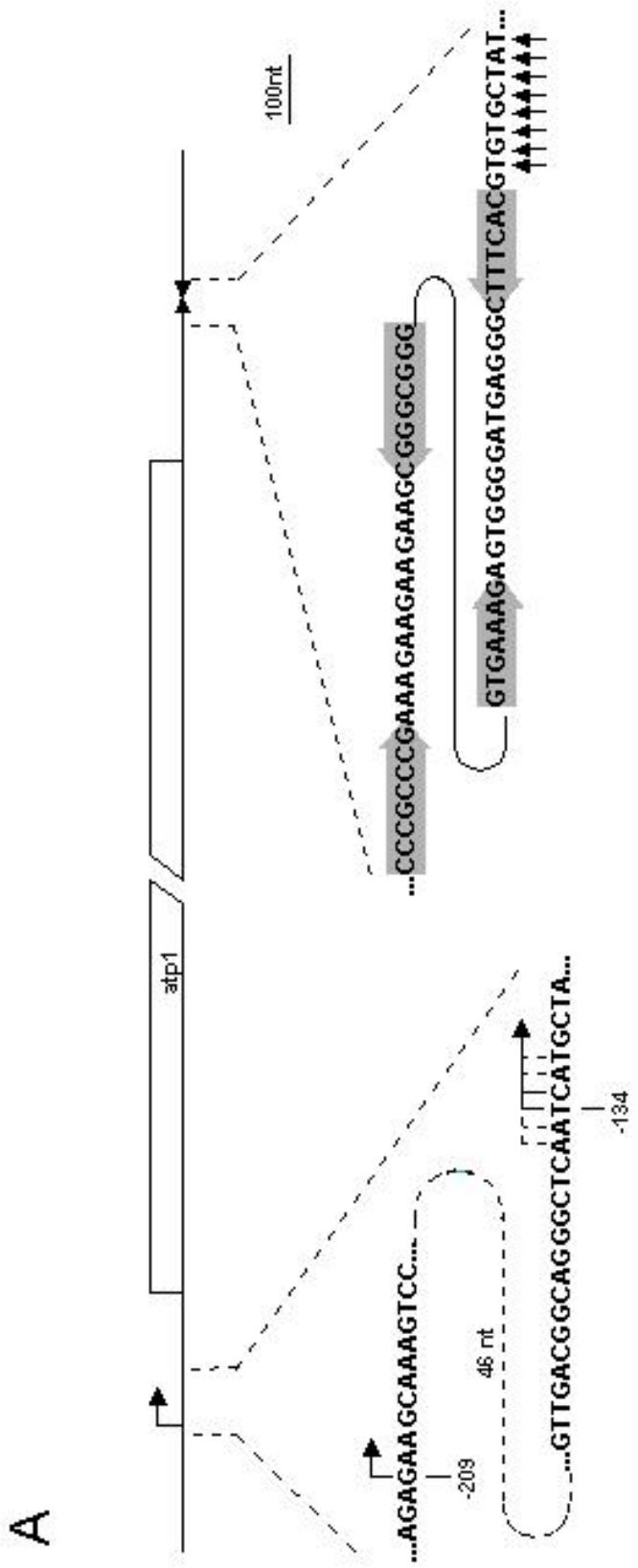
Σ 29

Kuhn and Binder, Fig. 2





Kuhn and Binder, Fig. 3



Kuhn and Binder, Fig. 4

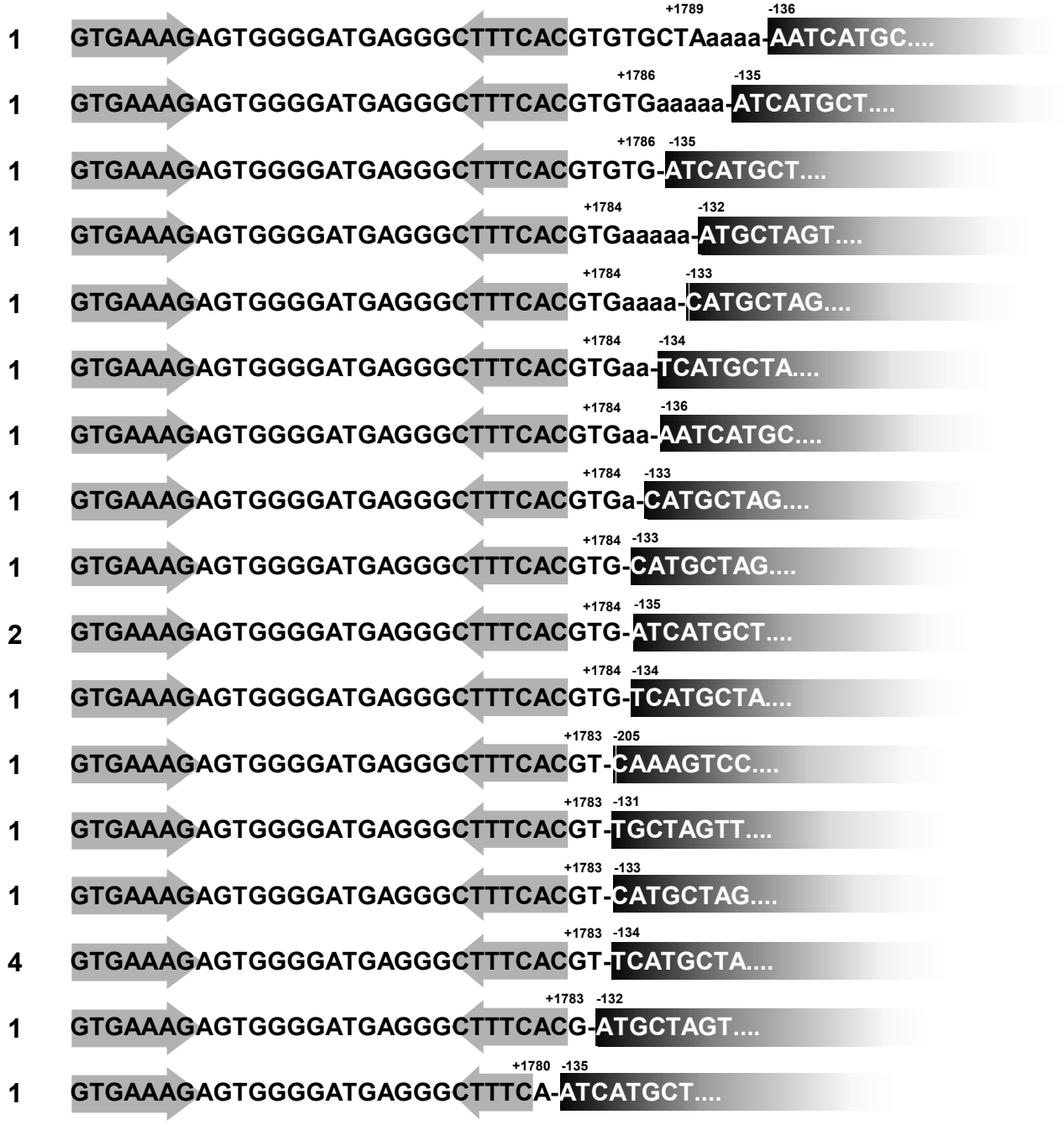
B



Σ 32

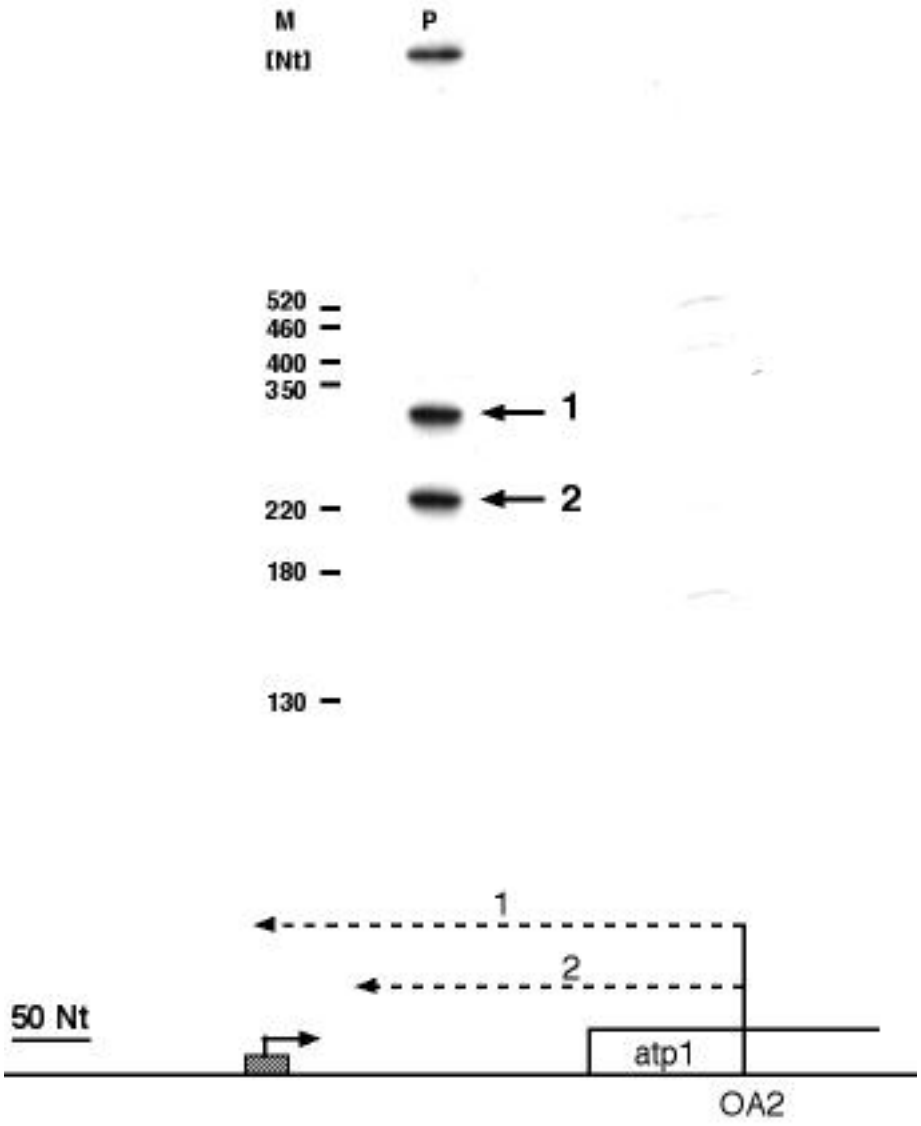
Kuhn and Binder, Fig. 4

C

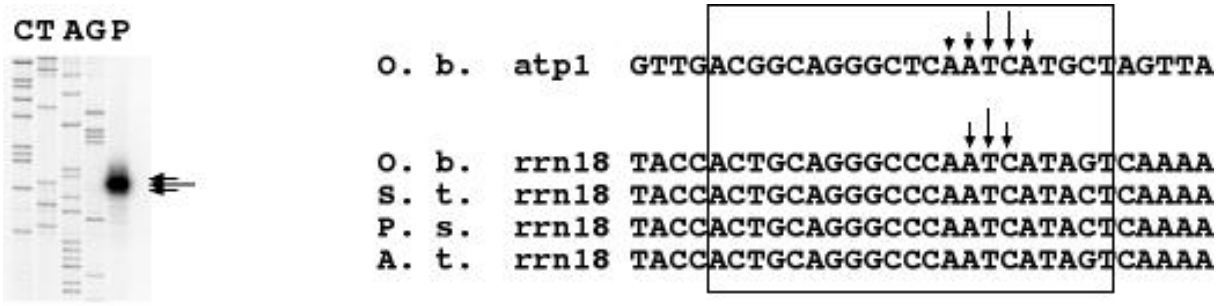


Σ 21

Kuhn and Binder, Fig. 4



Kuhn and Binder, Fig. 5



Kuhn and Binder, Fig. 6

6.1.4 RT-PCR analysis of 5' to 3'-end-ligated mRNAs identifies the extremities of *cox2* transcripts in pea mitochondria

© 2002 Oxford University Press

Nucleic Acids Research, 2002, Vol. 30, No. 2 439–446

RT–PCR analysis of 5′ to 3′-end-ligated mRNAs identifies the extremities of *cox2* transcripts in pea mitochondria

Josef Kuhn and Stefan Binder*

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

Received October 8, 2001; Revised and Accepted November 9, 2001

DDBJ/EMBL/GenBank accession no. AJ414385

ABSTRACT

Gene expression in plant mitochondria is still inadequately analyzed. To learn more about transcription and RNA processing in plant mitochondria, the 5′- and 3′-RNA extremities and the promoters of the cytochrome oxidase gene (*cox2*) were analyzed in pea. Both 5′ and 3′ ends of *cox2* transcripts were examined by RT–PCR across the ligation site of circularized mitochondrial RNA as template. This approach identified 5′ ends a few nucleotides shorter than three major 5′ ends mapped by primer extension analysis. Presumably, only monophosphate 5′ ends derived from processing can be ligated. *In vitro* transcription assays using a homologous mitochondrial protein extract from pea strongly suggest the major 5′ ends to derive from transcription initiation. The cDNA analysis of the head-to-tail ligated *cox2* mRNA identified 3′ ends within a thymidine stretch ~300 nt downstream of the reading frame in a sequence segment that was not present in the previous investigation of this gene. Nuclease S1 protection experiments confirmed this newly identified 3′ terminus and corroborated the validity of this technique in mRNA end analysis. The general use of the circularized RNA (CR)–RT–PCR approach for the simultaneous analysis of the 5′ and 3′ extremities of mRNA molecules is discussed.

INTRODUCTION

The mitochondrial DNAs (mtDNAs) of higher plants encode approximately 50–60 genes storing the genetic information for structural RNAs and proteins indispensable for the function of these organelles (1,2). To maintain and implement this genetic information a complete enzymatic equipment for DNA replication, transcription, RNA processing and translation within this compartment is required. Despite the progress that has been achieved in recent years, the majority of *cis* elements, *trans*-acting factors and mechanisms of plant mitochondrial gene expression are still unknown (3). For instance, the detailed analysis of transcription initiation sites in maize and pea revealed the promoter requirements, but the number of promoters analyzed or even definitely identified is still very limited and the overall picture of plant mitochondrial promoters remains patchy (4–8). Even less is known about the processes at the opposite end of the mRNA. Whereas transcription termination has not been investigated at all to date, our understanding of post-transcriptional processing of mitochondrial RNAs is likewise very limited. Beside RNA splicing and RNA editing altering the primary structure of an RNA molecule, various RNA processing steps in plant mitochondria promote or inhibit degradation of an RNA and are thus important determinants of RNA stability (3). This includes the alteration of the 3′ ends either by post-transcriptional trimming or the addition of non-encoded nucleotides (mostly adenosines) to the 3′ ends of at least some RNA species. This polyadenylation of transcripts was found in mitochondria from sunflower and maize and was suggested to accelerate degradation of these RNAs (9,10). Poly(A) tails were also found downstream of double stem–loop structures in *Oenothera atp1* and pea *atp9* transcripts. However, in the *atp9* mRNAs only the post-transcriptional 3′-end formation is substantially accelerated, whereas in contrast to mRNAs without stem–loop structure the complete degradation of the mRNA seems only slightly stimulated (11). Short extensions primarily composed of adenosines and cytidines were observed at the 3′ ends of pea *atp9* and maize *rps12*, *cox2* and *atp9* transcripts, but their function is unknown at present (11,12). The number of identified and thoroughly analyzed 3′-transcript termini, an absolute prerequisite for the investigation of both transcription termination events and mechanisms for post-transcriptional 3′-end formation, is still insufficient.

In the present work transcripts of the mitochondrially encoded gene *cox2* from pea are investigated. The circularized RNA (CR)–RT–PCR method employed simultaneously and rapidly identifies the 5′ and 3′ ends of the pea *cox2* mRNA. Controls with primer extension, nuclease S1 protection and *in vitro* transcription analyses calibrate and confirm the validity of this technique.

*To whom correspondence should be addressed. Tel: +49 731 502 2625; Fax: +49 731 502 2626; Email: stefan.binder@biologie.uni-ulm.de

440 Nucleic Acids Research, 2002, Vol. 30, No. 2

MATERIALS AND METHODS

RT–PCR analysis of circularized RNA

Total mitochondrial RNA was extracted from isolated and purified mitochondria as described by Binder *et al.* (13). Between 5 and 9 µg of this mtRNA was self-ligated with T4 RNA ligase in a total volume of 10 µl under conditions recommended by the manufacturer of the enzyme (Roche Molecular Biochemicals). After an adjustment of the total volume to 400 µl, proteins were removed by phenol–chloroform extraction following standard protocols and the total volume was reduced to 10 µl. First-strand synthesis was then carried out with primer PC2 (105 to 125) and Superscript reverse transcriptase on 5–9 µg ligated RNA following the instructions of the respective manual (Life Science Technologies). DNA fragments comprising the 5′–3′-ligated mRNA ends were subsequently amplified with primer pairs C4/PC8 (648–669) followed by a second PCR with nested primers PC3/PC9 (670–691) using 0.5–0.75 kb amplification products of the first PCR as template. The products were finally cloned into an *Xcm*I linearized pBluescript vector according to standard procedures followed by sequence analysis of 15 different *cox2* cDNA clones (14). The locations of the primers are given in respect to the translation initiation codon (+1).

Nucleic acids analysis

Primer extension analysis was carried out with primer PC3 (–201 to –218) as described by Marchfelder *et al.* (15). The 5′ ends were identified by electrophoresis of the primer extension products alongside sequencing fragments obtained with the same primer (PC3). Nuclease S1 protection experiments were essentially done as described by Hoffman *et al.* (16) with the following modifications: a 0.5 kb *Hin*dIII–*Eco*RI fragment covering nucleotides 200–700 downstream of the reading frame was labeled with [α-^{32}P]dATP by the Klenow fragment of DNA polymerase I and used as antisense probe.

In vitro transcription analysis

Pea mitochondria were isolated, purified and processed to transcriptionally active lysates as described by Binder *et al.* (7). To generate the pea *cox2* genomic clone pcox2HE75 a respective DNA fragment was amplified by PCR with primers PC1 (–628 to –608) and PC2 digested with *Hin*dIII–*Eco*RV and cloned into pBluescript vectors digested with the same enzymes. Template pcox2HC468 was obtained after amplification with primers PC1 and C4 (48 to 31), digestion of the fragment with *Hin*dIII and cloning into a *Hin*dIII–*Sma*I digested pBluescript vector. PCR was performed using KlenTaq polymerase according to the conditions suggested by the manufacturer with pea mtDNA as template. Promoter activity was tested with ~100 ng of the DNA templates linearized with *Nco*I (pcox2HE75) and *Bst*XI (pcox2HC468) as described by Binder *et al.* (7).

Miscellaneous methods

Filters (Hybond N) carrying *Hin*dIII and *Eco*RI pea mtDNA libraries were prepared and hybridized according to a protocol given by the manufacturer (Amersham Pharmacia Biotech). Oligonucleotide pcox2INS-1 (1063–1045) was used as probe for the isolation of clones representing the *cox2* gene and the respective genomic downstream region.

Sequence analyses of the genomic and cDNA clones including those used as DNA templates for *in vitro* transcription were done with a T7 sequencing kit and [^{35}S]dATP or with a Thermo Sequenase fluorescent labeling kit according to the instructions of the manufacturer (Amersham Pharmacia Biotech). Sequencing products were either separated by conventional polyacrylamide gel electrophoresis or on an ALF express sequencer.

RESULTS

An RT–PCR analysis of 5′–3′ self-ligated *cox2* mRNA reveals a new 3′ end within a oligothymidine stretch

A previous analysis of the pea *cox2* gene and its expression had revealed this gene to be transcribed into monocistronic transcripts, with two 5′ ends located at nucleotide positions –302 and –285 upstream of the reading frame and a single 3′ end 193–195 nt downstream of the translation stop codon (17). Thus, this gene represents a good candidate to test a method employing RT–PCR across the 5′–3′-end-ligated mRNAs, which would allow to simultaneously investigate both extremities of the transcripts and which has successfully been used for the analysis of land snail mitochondrial tRNAs (18). To this end, total RNAs from pea mitochondria were self-ligated with T4 RNA ligase and specific first-strand cDNA synthesis was initiated with primer PC2 followed by a PCR with primers PC8 and C4 (Fig. 1). Separation of the amplification products revealed three major molecule species of ~650, 420 and 170 bp, and a smaller one most likely representing non-incorporated primers (Fig. 1, bottom part). A PCR product of ~650–670 nt was expected if the 5′ and 3′ termini of the steady state 1.2 kb transcripts are correctly ligated. Accordingly, DNA fragments with sizes between 0.5 and 0.75 kb were recovered from the gel and used as templates in a second PCR with nested primers PC3/PC9 (Fig. 1). This generated a predominant cDNA fragment of ~500 nt, slightly larger than the 450 nt expected for ligation at the previously reported mRNA extremities (17). To cover these sizes cDNA fragments between ~420 and 530 nt were cloned and 15 clones were sequenced. No PCR fragment could be obtained by the same experimental strategy without RNA ligase indicating that 5′–3′-end-ligated *cox2* mRNAs are not detectable in the steady state RNA, which, however, does not necessarily exclude the existence of RNA molecules with such a structure.

The 5′ end in a single clone maps 6 nt upstream of position –285, whereas the 5′ termini of the residual 14 clones map 5–24 nt downstream of this position (Fig. 2). In contrast to this congruent assignment of the 5′ sequences, the 3′-terminal nucleotides do not match the previously published sequence (17). The cDNA sequences could be aligned with genomic sequences at primer PC9 from nucleotide 109 upstream of the translation stop codon to 237 nt downstream of the reading frame, ~44 nt beyond the originally mapped 3′ end (+193 to +195). This sequence is followed by 61–64 nt, depending on the clone, not found in the published genomic sequence (17). To test whether the newly identified sequence segment is present at the 5′ or the 3′ ends, corresponding RT–PCR analyses were carried out. Whereas no cDNA fragment was observed with an RT–PCR strategy consistent with an attachment of the newly identified sequence at the 5′ end of *cox2* transcripts,

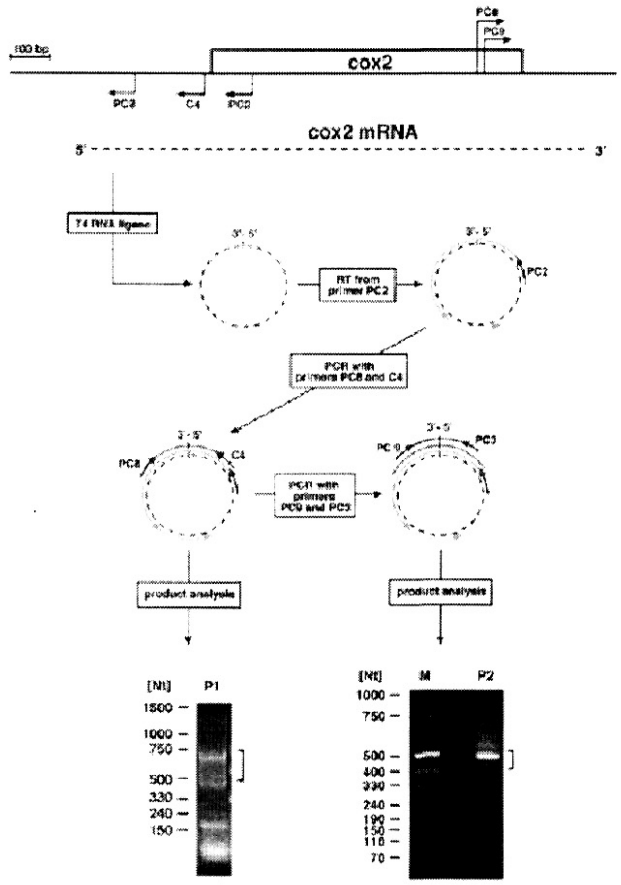


Figure 1. CR–RT–PCR analysis of *cox2* transcripts. Location and orientation of the primers (bent arrows) used on the *cox2* transcripts (dashed line) are indicated in a cartoon of the pea *cox2* locus (upper part). The 5′ and 3′ ends of the transcripts are joined using T4 RNA ligase. The circular form of the RNA is then used as template for first-strand synthesis initiated from primer PC2 using reverse transcriptase (RT) followed by two consecutive amplification reactions with primer pairs C4/PC8 and PC3/PC9, each enclosing the ligation site. Products obtained in the CR–RT–PCR analysis are separated on agarose gels (bottom part). Sizes of coelectrophoresed marker molecules are given in nucleotides. Angular brackets indicate the range of excised fragments.



Figure 2. 5′ and 3′ ends detected by the CR–RT–PCR. Sequences of the cDNA covering the 5′–3′-ligation sites of pea *cox2* transcripts. The numbering of the 5′ end refers to the translation start codon and the 3′ numbers are relative to the stop codon. The numbers on the left indicate the frequency with which the distinct 5′ and 3′ ends are found in individual clones. The ligation site is indicated by a horizontal dash. Sequences corresponding to the 5′-terminal part are given in white letters in fading black boxes.

cDNA fragments were generated in two amplification reactions corresponding to an mRNA containing the new sequence at the 3′ end. This suggests these nucleotides to be an integral part of the 3′-non-translated region of the *cox2* mRNA (data not shown).

Mapping three major 5′-transcript termini upstream of the pea *cox2* gene

Since the 5′ ends detected in the CR–RT–PCR analysis did not match the previously mapped termini these were re-examined by primer extension analysis from primer PC3, which anneals to sequences 201–218 nt upstream of the translation start codon. This analysis revealed three major ends in repeated experiments in the initially described region, but all differing from the previously mapped sites, and also from the 5′ ends detected in the cDNA derived from the CR–RT–PCR analysis (Figs 2 and 3). The 5′ ends (sites 1–3) were found at positions –316, –296 and –279 with respect to the ATG.

Two of the newly identified 5′ ends (–296 and –279) are associated with sequences showing strong similarity with the consensus promoter motif identified in dicot plant species (8). To test the functional significance of this similarity, the entire region was further analyzed.

In vitro transcription analysis confirms promoters upstream of the *cox2* reading frame

In order to investigate whether the identified ends indeed derive from transcription initiation, two different DNA templates comprising the detected 5′ ends were incubated with a transcriptionally active mitochondrial protein lysate. The sizes of the strongest signals (Fig. 4, signals b and e in lanes A and B) correlate well with *de novo* transcription initiation at site 2 (corresponding to position –296), strongly suggesting the surrounding sequences to function as a bona fide mitochondrial promoter. Signals of slightly lower intensities correspond to transcription initiation events at site 3 (–279), supporting the respective 5′ end to be derived from a transcription start site (Fig. 4, signals c and f in lanes A and B). Two products with much lower abundance correspond to run-off transcripts expected upon transcription initiation at site 1 (–317). This suggests low but active transcription initiation at this site, even though the surrounding sequences do not resemble any known promoter sequence from dicot plant mitochondria (Fig. 4, signals a and d in lanes A and B).

The new sequence segment is present in the genomic DNA downstream of the *cox2* gene

Although the presence of 3′-end non-encoded nucleotides of different identities has been reported for several plant mitochondrial transcripts, the attachment of non-homopolymeric sequences of a length of >60 nt as found here (see above)

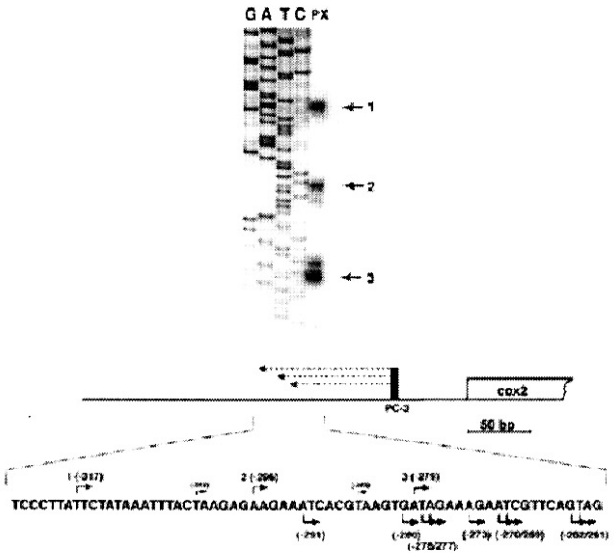


Figure 3. Primer extension analysis of pea *cox2* transcripts. Three major 5′-transcript termini (indicated by numbered arrows on the right) were detected by running the extension products (lane PX, illustrated as dashed lines in the bottom part) alongside a sequencing ladder (lanes G, A, T, C). Both primer extension and sequencing analysis were carried out with primer PC3 (vertical black bar) annealing to the mRNA upstream of the *cox2* reading frame (open box). The sequence surrounding the detected 5′ ends (numbered bent arrows above the sequence) is shown in the bottom part. Nucleotide positions relative to the ATG are given in brackets. Previously mapped 5′ termini are also given above the sequence (17). 5′ Extremities found by the CR–RT–PCR analysis are given below. All numberings are in respect to the translation start codon (+1).

seems highly unlikely (11,12). Therefore, we examined the 3′ region of the pea *cox2* gene on the level of the genomic DNA. Using oligonucleotide p*cox2*INS-1, corresponding to part of the newly identified sequence, a 1.8 kb *Eco*RI and a 1.7 kb *Hin*dIII clone were isolated from pea mtDNA libraries. Sequence analysis of the former showed this to carry the complete *cox2* reading frame and ~350 nt as well as 650 nt 5′- and 3′-flanking sequences, respectively (accession no. AJ414385). This clone corresponds to the *Eco*RI clone used for the previously published sequence analysis (17). The *Hin*dIII clone represents the *cox2* 3′ region and overlaps with the *Eco*RI clone by 600 nt. This clone also contains the newly identified sequence confirming its presence in the genomic sequence (Fig. 5). Thus, the CR–RT–PCR of *cox2* transcripts locates the 3′-transcript terminus ~100 nt downstream of the previously mapped end (17). To confirm this by an independent experimental approach, an S1 protection analysis covering the 3′-non-translated region 200–700 nt downstream of the *cox2* reading frame was performed (Fig. 5). The protected fragment of ~100 nt corresponds to a 3′ end ~300 nt downstream of the translation stop codon, consistent with the 3′-transcript termini detected in the CR–RT–PCR clones (Fig. 6). These results confirm the validity of the CR–RT–PCR method for 3′-end mapping of mitochondrial transcripts. Comparison of the cDNA sequences derived from the CR–RT–PCR analysis with genomic sequences revealed no non-encoded nucleotides at either end of the *cox2* mRNA.

DISCUSSION

Pea *cox2* mitochondrial transcripts: test object for the CR–RT–PCR analysis of mRNA extremities

Previous analysis of the *cox2* gene encoded in pea mitochondria had indicated a relatively simple transcript pattern, which suggested this gene as an ideal candidate for a detailed functional investigation of its expression. Two 5′-transcript termini had been mapped in a region containing sequence elements with substantial similarity to dicot plant mitochondrial promoters, and a nuclease S1 protection analysis suggested a 3′ end ~193–195 nt downstream of the reading frame (17). Later analysis of cDNAs originating from this gene indicated several nucleotide differences compared with genomic sequences recommending a detailed descriptive re-examination (19).

Re-examination of the *cox2* gene encoded in pea mitochondria

We started the re-investigation of the transcript extremities followed by a complete sequence analysis of the gene and its 5′ and 3′ genomic environments. The latter revealed several nucleotide differences with respect to the previously published sequences with the additional 54 nt in the 3′-terminal part of the mRNA as the most striking discrepancy (see accession no. AJ414385 and Fig. 5). All non-RNA editing DNA/RNA differences between the previously published genomic (17) and the cDNA sequences (19) were found to be correct in the latter. From the 14 RNA editing DNA/RNA differences identified by the cDNA analysis, 12 were confirmed in our genomic sequence, whereas the editing events at 44 nt upstream and 412 nt downstream of the ATG indicated by C/T differences have to be corrected. At these sites thymidines are already present in our genomic pea sequence (19). Comparison of the new cDNA sequences obtained from the CR–RT–PCR covering the complete 3′- and partially the 5′-non-translated parts of the mRNA does not indicate any further editing site.

Similarity between the pea *cox2* promoter sequences and conserved nonanucleotide promoters from dicot plants

In contrast to the two *cox2* mRNA 5′ ends previously mapped by nuclease S1 protection experiments, the primer extension technique applied here revealed three 5′ ends each differing by several nucleotides from the previously mapped ends. These discrepancies are probably caused by the different techniques used, an often observed phenomenon. Comparison of the sequences at sites 2 and 3 shows substantial similarities to the pea *atp9* promoter, which has been investigated by detailed functional *in vitro* transcription studies (7,8). As shown in Figure 7, well conserved nonanucleotide motifs (CNM) are present at both sites 2 (–296) and 3 (–279) with minor variations at the first two positions of this motif at site 2. In addition, the presence of upstream-located AT-rich regions and the presence of at least one purine at positions +3 and +4 fulfill the requirements of a typical dicot plant mitochondrial promoter (8). Consequently, run-off transcripts consistent with transcription initiation at these sites are detected in the functional *in vitro* studies confirming the function of these sequences as mitochondrial promoters. In contrast, a completely different sequence surrounds the most upstream 5′ end at site 1 (–317), with no similarity to either CNM or to other non-CNM-type

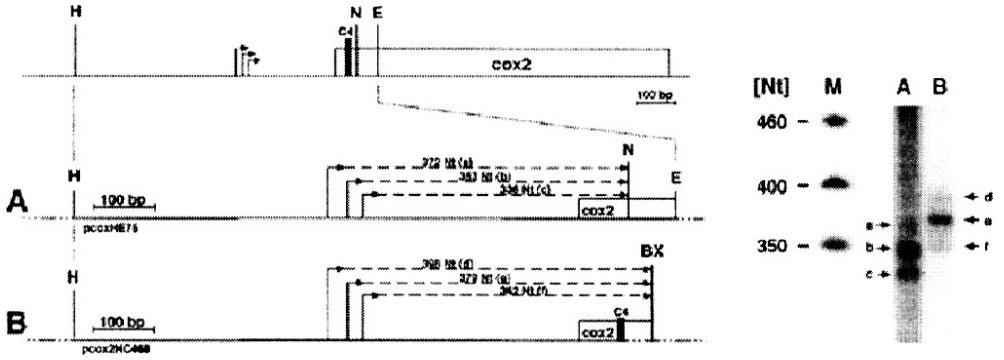


Figure 4. *In vitro* transcription analysis of pea *cox2* DNA templates. A diagram of the pea *cox2* genomic region and the respective DNA templates used for *in vitro* transcription analysis are shown in the left part. Both DNA templates (**A** and **B**) contain the 5′ region of the *cox2* reading frame (indicated by the open box) including the three 5′ ends (bent arrows) mapped as detailed above. Expected sizes of the run-off products (dashed lines) of the *Nco*I and *Bst*XI linearized templates (A) pcox2HE75 and (B) pcox2HC468, respectively, are indicated. Restriction sites are given for *Hin*dIII (H), *Eco*RV (E) and *Nco*I (N). The *Bst*XI site (BX) used for the linearization of pcox2HC468 is located in vector sequences (lines interrupted by dots). The location of primer C4 determining the downstream end of the cloned mitochondrial sequences is indicated (C4, vertical black bar). The *in vitro* transcription analysis is shown in the right-hand part. The designation of the lanes corresponds to the templates used in the *in vitro* assays. The numbering of the detected *in vitro* run-off products (a–g, indicated by numbered arrows) corresponds to the one given for the expected products.



Figure 5. Sequence context of the pea *cox2* 3′ region. Two pea mtDNA clones designated E21 (*Eco*RI 1.8 kb) and H1 (*Hin*dIII, 1.7 kb) (indicated by solid lines) representing the *cox2* genomic environment were used for the genomic sequence analysis of the respective *cox2* region. The new sequence segment identified in the CR–RT–PCR analysis is highlighted by a gray box. The previously mapped (1) and the newly identified 3′-transcript termini (2) are indicated by vertical arrows above the sequence. Restriction sites are given by dashed vertical lines for *Eco*RI (E) and *Hin*dIII (H). The probe used in the nuclease S1 protection assays (Fig. 6) is given as a hatched bar. Position of the radioactive label is indicated by a white P in a black circle.



Figure 6. Nuclease S1 protection analysis confirms the new mRNA 3′ end detected by the CR–RT–PCR. Protected products (lane 5) were coelectrophoresed with products obtained in control experiments and with DNA marker fragments (lanes M), whose sizes are given in the left margin. The control reactions are the antisense probe without any further treatments (lane 1), incubated under S1 conditions without nuclease and RNA (lane 2), incubated with 75 U S1 and without RNA (lane 3) or 40 µg of wheat germ tRNA products (lane 4). The size of the protected fragment of ~100 nt corresponds to a 3′ end ~300 bp downstream of the translation stop codon, which is consistent with the 3′ terminus mapped by the CR–RT–PCR analysis.

transcription initiation sites identified in mitochondria of various dicot plant species (20–23). Nevertheless, weak signals that could correspond to run-off products initiated at this site are detected in the *in vitro* analysis, suggesting promoter activity of this sequence, the first observation of a non-CNM promoter activity *in vitro*.

Primary and secondary structures at the 3′ end of the pea *cox2* transcripts

At present little is known about the generation of 3′ termini in higher plant mitochondria. Generally, 3′-end formation can be achieved either directly by transcription termination or by subsequent post-transcriptional RNA processing. The modes and structures involved in the former are completely unknown, and the double stem–loop structures located immediately upstream of several characterized 3′ ends are not able to terminate transcription *in vitro*. These structures are instead processing signals impeding the progression of 3′–5′ exonucleolytic degradation in post-transcriptional 3′ trimming (24,25). How the *cox2* 3′ mRNA ends investigated here in pea mitochondria are formed remains unclear at present. Two alternative stem–loop structures can be folded a few nucleotides upstream of the 3′ termini. These structures and the location of the 3′ ends within a polythymidine stretch of 9 nt resemble prokaryotic transcription terminators and may suggest this mRNA terminus to be derived by transcription termination (Fig. 8) (26). Thus, detailed *in vitro* studies are required to learn more about 3′-end formation of pea *cox2* transcripts and plant mitochondrial 3′-end formation in general.

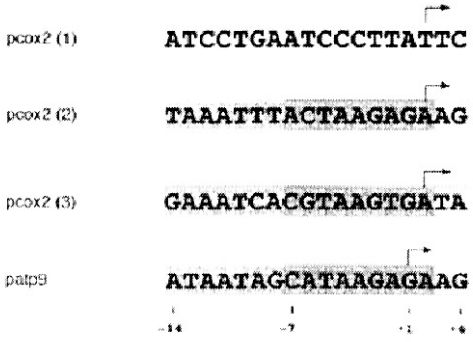


Figure 7. Sequences of pea *cox2* promoters. The sequences surrounding the three different 5′ ends of pea *cox2* transcripts [pcox2 (1–3)] are aligned with the −14 to +4 promoter sequence located upstream of the pea *atp9* (patp9) gene. Whereas sequences at sites 2 and 3 were aligned for best nucleotide consistence with the pea *atp9* promoter, the sequence at site 1 is aligned with respect to the detected 5′ ends (indicated by bent arrows). Nucleotide positions are given for the pea *atp9* promoter referring to the transcription start site (+1). The AT-rich region (AT-Box, −14 to −8) and the CNM (−7 to +2) as well as purine-rich nucleotide positions +3/+4, all hallmarks of the CNM-type dicot plant mitochondrial promoters, are boxed and highlighted.

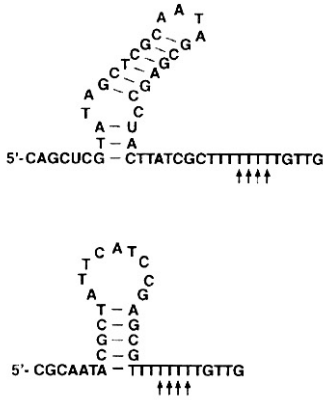


Figure 8. Potential secondary structures at the 3′ end of the pea *cox2* transcripts identified here. The 3′ ends detected by nuclease S1 and CR–RT–PCR analysis are indicated by vertical arrows beneath the sequence.

Non-encoded nucleotides at the 3′ end of pea *cox2* transcripts

Besides the mapping of RNA 5′ and 3′ extremities the CR–RT–PCR analysis further allows the simultaneous detection of non-encoded nucleotides. However, in the case of pea *cox2* transcripts non-encoded nucleotides could not be detected by this approach. Although the total number of 15 clones analyzed might be too low, an alternative strategy using a 3′-RACE experiment initiated with oligo(dT) adapter primer revealed only a single clone deriving from an annealing of the oligo(dT) primer to an endogenous oligoadenosine stretch, but failed to detect 3′-non-encoded nucleotides (A.Dospil, J.Kuhn and S.Binder, unpublished data). Since this latter approach has been successfully used in the analysis of other mitochondrial transcripts (9–11), the presence of 3′-non-encoded nucleotides at pea *cox2* transcripts is highly unlikely.

CR–RT–PCR, a convenient method for the investigation of 5′ and 3′ extremities of plant mitochondrial RNA

The mapping of 5′ ends by an RT–PCR analysis of circularized mRNA revealed different termini, most of them located up to 20 nt downstream of site 3. Whereas one clone ends between sites 2 and 3, the 5′ ends of three clones coincide with a guanidine at position –280, 1 nt upstream of the major product at site 3 at position –279. This distribution of the 5′ ends is in contrast to the primer extension experiments, which clearly indicate three major 5′ ends. The exact reason for this discrepancy is unclear at present. We assume that 5′ ends with di- or triphosphate groups present at primary transcripts whose 5′ ends derive from *de novo* initiation events, are ligated with much lower efficiency (or not at all) compared with termini with 5′-monophosphate groups at transcripts originating from processing events. This bias is strongly suggested by the mechanism of the ATP-dependent T4 RNA ligase reaction. In the first step, AMP, which is formed by the release of pyrophosphate from ATP, is bound to the 5′-phosphoryl-terminated nucleic acid donor via a pyrophosphate bond yielding an adenylated donor [A5′pp-5′Np(Np)$_x$]. In the next step, AMP is released from the adenylated donor and replaced by a phosphodiester bond between the donor and a 3′-hydroxyl-terminated acceptor. The latter ligation step, which is independent from ATP, can be performed by T4 RNA ligase with various adenylated donor molecules and 3′ hydroxyl acceptor, but not with an A5′ppp5′G, which would theoretically result from the first reaction step from a 5′ diphosphate donor molecule (27,28). Concluding, it is thus highly unlikely that the 5′ di- or triphosphate ends found at 5′ termini from transcription initiation events can be accepted as substrates of the RNA ligase reactions.

The *cox2* mRNAs slightly shortened at the 5′ end are possibly products of processing or degradation events. These transcripts are most likely shortened at the 5′ end by endo- or exonucleolytic activities, rather than resulting from initiation events downstream of the major initiation sites and thus carry monophosphate groups at the 5′ ends. This hypothesis is further supported by the sequence analysis of 21 cDNA clones obtained from CR–RT–PCR investigation of self-ligated *Oenothera atp1* transcripts, which revealed no 5′ end identical with the originally mapped primary 5′ end (J.Kuhn and S.Binder, manuscript in preparation). Instead, an alternative 5′ terminus was found in 20 of these clones, which is located ~76–79 nt downstream of the transcription initiation site as confirmed by primer extension experiments and which is most likely derived from a processing event. Only a single clone contained a 5′ end close to the initiation site, but is still lacking the terminal 5 nt.

If the RNA ligase indeed prefers monophosphate over di- or triphosphate 5′ ends, which has of course to be investigated experimentally, the CR–RT–PCR method will give a straightforward assignment of 5′ ends originating from processing or degradation events without any pretreatment. On the other hand, primary 5′ termini could only be ligated once di- or triphosphate groups have been converted to monophosphate groups. Such a pretreatment, e.g. by consecutive phosphatase and kinase reactions, would also make these termini accessible for the RNA ligase. Such a differential pretreatment of the

RNA could allow the assignment of the origin of any 5′ end from either *de novo* transcription initiation or from processing events.

General suitability of the method for mapping of 5′- and 3′-RNA extremities

The application of the CR–RT–PCR for the transcript analysis of the pea *cox2* as well as pea *atp9* and *Oenothera atp1* genes (J.Kuhn and S.Binder, manuscript in preparation) revealed the 5′ and 3′ extremities mapped by this technique to be generally consistent with respective ends determined by other methods. For instance the *cox2* 3′-transcript termini detected here in the CR–RT–PCR analysis could be clearly confirmed by a nuclease S1 experiment (Fig. 6). The 3′ ends mapped previously most likely originate from a misleading nuclease S1 protection strategy using a fragment that is entirely complementary to internal sequences of the mRNA and should thus have been entirely protected. The observed shortening by a few nucleotides is perhaps caused by the use of too stringent nuclease S1 digestion conditions.

The application of the CR–RT–PCR technique presented here for the analysis of the above mentioned transcripts is facilitated by the knowledge of both ends, which allowed to set up the appropriate PCR strategy and to preselect cDNA fragments obtained in the amplification reactions. But will this technique also be suitable for 5′- and 3′-end mapping of totally uncharacterized transcripts? This will probably be rather difficult and at least the rough size of the transcript(s), e.g. determined by northern blot experiments, will help to estimate the extent of the 5′- and 3′-untranslated regions. This in turn will allow a prediction of the size(s) of the expected RT–PCR fragment(s). Such a northern blot analysis will also reveal the overall transcript pattern, which for plant mitochondrial genes can be very complex and will complicate the 5′- and 3′-end mapping by any method.

The requirement of monophosphate 5′ termini defines the types of RNA that can be investigated by CR–RT–PCR. The analysis of RNA species such as those derived from processed mRNAs in any system like mitochondria, plastids or bacteria will be straightforward. For capped RNAs this method has previously been successfully used to characterize the 5′- and 3′-terminal regions of non-polyadenylated linear RNA genomes from several tick-borne flaviviruses. In this case, the 5′-terminal cap had to be removed by tobacco acid pyrophosphatase to allow the 5′–3′-end-ligation by T4 RNA ligase (29).

However, the same strategy should also allow the analysis of polyadenylated mRNAs, whose 5′ ends are presently often mapped by ligation of an anchor primer to the 5′ end instead of circularization of the mRNA (30).

ACKNOWLEDGEMENTS

We thank Bärbel Weber for excellent technical assistance. This work was supported by grants from the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie to S.B. and a fellowship of the Studienstiftung des Deutschen Volkes to J.K.

REFERENCES

1. Unseld,M., Marienfeld,J.R., Brandt,P. and Brennicke,A. (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nature Genet.*, **15**, 57–61.
2. 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.
3. Hoffmann,M., Kuhn,J., Däschner,K. and Binder,S. (2001) The RNA world of plant mitochondria. In Moldave,K. (ed), *Progress in Nucleic Acid Research and Molecular Biology*. Academic Press, San Diego, Vol. 70, pp. 119–154.
4. Rapp,W.D. and Stern,D.B. (1992) A conserved 11 nucleotide sequence contains an essential promoter element of the maize mitochondrial *atp1* gene. *EMBO J.*, **11**, 1065–1073.
5. Rapp,W.D., Lupold,D.S., Mack,S. and Stern,D.B. (1993) Architecture of the maize mitochondrial *atp1* promoter as determined by linker-scanning and point mutagenesis. *Mol. Cell. Biol.*, **13**, 7232–7238.
6. Caoile,A.G. and Stern,D.B. (1997) A conserved core element is functionally important for maize mitochondrial promoter activity *in vitro*. *Nucleic Acids Res.*, **25**, 4055–4060.
7. 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.
8. Dombrowski,S., Hoffmann,M., Guha,C. and Binder,S. (1999) Continuous primary sequence requirements in the 18-nucleotide promoter of dicot plant mitochondria. *J. Biol. Chem.*, **274**, 10094–10099.
9. 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.
10. 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.
11. 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.
12. 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.
13. Binder,S., Marchfelder,A. and Brennicke,A. (1994) RNA editing of $tRNA^{Phe}$ and $tRNA^{Cys}$ in mitochondria of *Oenothera berteriana* is initiated in precursor molecules. *Mol. Gen. Genet.*, **244**, 67–74.
14. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Habor Laboratory Press, Cold Spring Harbor, NY.
15. Marchfelder,A., Brennicke,A. and Binder,S. (1996) RNA editing is required for efficient excision of $tRNA^{Phe}$ from precursors in plant mitochondria. *J. Biol. Chem.*, **271**, 1898–1903.
16. Hoffmann,M., Dombrowski,S., Guha,C. and Binder,S. (1999) Cotranscription of the *rpl5–rps14–cob* gene cluster in pea mitochondria. *Mol. Gen. Genet.*, **261**, 537–545.
17. Moon,E., Kao,T.H. and Wu,R. (1985) Pea cytochrome oxidase subunit II gene has no intron and generates two mRNA transcripts with different 5′-termini. *Nucleic Acids Res.*, **13**, 3195–3212.
18. Yokobori,S. and Paabo,S. (1995) Transfer RNA editing in land snail mitochondria. *Proc. Natl Acad. Sci. USA*, **92**, 10432–10435.
19. Covello,P.S. and Gray,M.W. (1990) Differences in editing at homologous sites in messenger RNAs from angiosperm mitochondria. *Nucleic Acids Res.*, **18**, 5189–5196.
20. 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.
21. Fey,J. and Marechal-Drouard,L. (1999) Expression of the two chloroplast-like $tRNA^{Asn}$ genes in potato mitochondria: mapping of transcription initiation sites present in the *trnN1–trnY–nad2* cluster and upstream of *trnN2*. *Curr. Genet.*, **36**, 49–54.
22. Remacle,C. and Marechal-Drouard,L. (1996) Characterization of the potato mitochondrial transcription unit containing 'native' *trnS* (GCU), *trnF* (GAA) and *trnP* (UGG). *Plant Mol. Biol.*, **30**, 553–563.

23. Binder,S., Thalheim,C. and Brennicke,A. (1994) Transcription of potato mitochondrial 26S rRNA is initiated at its mature 5′ end. *Curr. Genet.*, **26**, 519–523.
24. 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.
25. 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.
26. Lesnik,E.A., Sampath,R., Levene,H.B., Henderson,T.J., McNeil,J.A. and Ecker,D.J. (2001) Prediction of rho-independent transcriptional terminators in *Escherichia coli*. *Nucleic Acids Res.*, **29**, 3583–3594.
27. Sugino,A., Snopek,T.J. and Cozzarelli,N.R. (1977) Bacteriophage T4 RNA ligase. Reaction intermediates and interaction of substrates. *J. Biol. Chem.*, **252**, 1732–1738.
28. England,T.E., Gumport,R.I. and Uhlenbeck,O.C. (1977) Dinucleoside pyrophosphate are substrates for T4-induced RNA ligase. *Proc. Natl Acad. Sci. USA*, **74**, 4839–4842.
29. Mandl,C.W., Heinz,F.X., Puchhammer-Stockl,E. and Kunz,C. (1991) Sequencing the termini of capped viral RNA by 5′–3′ ligation and PCR. *Biotechniques*, **10**, 484–486.
30. Fromont-Racine,M., Bertrand,E., Pictet,R. and Grange,T. (1993) A highly sensitive method for mapping the 5′ termini of mRNAs. *Nucleic Acids Res.*, **21**, 1683–1684.

6.2 Curriculum vitae

Name	Josef Martin Kuhn
Anschrift	Oberer Hasenkopfweg 54
	89075 Ulm
Geburtsdatum	28.05.1972
Geburtsort	Marktoberdorf, Deutschland
Nationalität	Deutsch
Familienstand	Ledig

Schulbildung

09/1982 - 07/1986	Carl-von-Linde Gymnasium Kempten, wohnhaft im bischöfliches. Studienseminar St. Magnus
09/1986 - 07/1992	Staatliches Gymnasium Kaufbeuren Abschluss: Abitur, "gut" (2,0)

Zivildienst

07/1992 - 09/1993	Kreiskrankenhaus Marktoberdorf: (Organisationsdienst, Mitarbeit hämatologisches Labor).

Hochschulstudium

10/1993 - 09/1998	Studiengang Biologie/Chemie Lehramt und Biologie Diplom, Universität Ulm
01/1998 - 09/1998	Diplomarbeit: „Funktionsanalyse von 3' Inverted-Repeats in Mitochondrien höherer Pflanzen“, Abt. Allg. Botanik, Universität Ulm Abschluss: Diplom-Biologe, „mit Auszeichnung“ (1,0)

Promotion

10/1998 - 5/2002	Abteilung Molekulare Botanik, Universität Ulm. Dissertation: „Funktionsanalyse der 3' mRNA Prozessierung in Mitochondrien höherer Pflanzen“

Stipendium

07/1999 - 3/2002	Promotionsstipendium der Studienstiftung des Deutschen Volkes

Tätigkeiten

03/1996	Mitarbeit, Abteilung Humangenetik, Universität Ulm
01/1997 - 03/1997	Studentische Hilfskraft, Abt. für vergleichende Neurobiologie, Universität Ulm
01/1998 - 10/1998	Studentische Hilfskraft, Abt. Allg. Botanik, Universität Ulm
10/1998-	Wissenschaftlicher Mitarbeiter: Betreuung von Literaturseminaren, Praktika, Diplomanden und Doktoranden, Begutachtung von Publikationen für Fachjournale

Zusatzqualifikationen	Betriebswirtschaftslehre (2 Semester) Pädagogische Grundkenntnisse PC-Administration
Sprachkenntnisse	Englisch verhandlungssicher Französisch gute Kenntnisse Spanisch Grundkenntnisse

6.3 Publications

Dombrowski S., M. Hoffmann, **J. Kuhn**, A. Brennicke and S. Binder (1998) On mitochondrial promoters in *Arabidopsis thaliana* and other flowering plants. In: Plant mitochondria: From Gene to Function. Backhuys Publishers, Leiden, Netherlands, 165-170

Gagliardi D., **J. Kuhn**, U. Spadinger, A. Brennicke, C. J. Leaver and S. Binder (1999) An RNA helicase (AtSUV3) is present in *Arabidopsis thaliana* mitochondria. FEBS Lett. 458: 337-342

Kuhn J., U. Tengler and S. Binder (2001) Transcript lifetime is balanced between stabilizing stem-loop structures and degradation-promoting polyadenylation in plant mitochondria. Mol. Cell. Biol., 21: 731-742

Hoffmann M., **J. Kuhn**, K. Däschner and S. Binder (2001) The RNA world of plant mitochondria. Prog. Nucleic Acid Res. Mol. Biol., 70: 119-154

Kuhn J. and S. Binder (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.

Kuhn J. and S. Binder (2002) 5' meets 3': head to tail linked mRNAs in plant mitochondria. Manuscript submitted.

Kuhn J. and Binder S. (2000) In plant mitochondria adenosines and cytidines are posttranscriptionally added to 3' ends of mRNAs. Abstract, SV 02-03, Botanikertagung 2000, Jena, Germany.

Kuhn J. and Binder S. (2000) In plant mitochondria 3' inverted repeats mediate the balance between transcript stability and degradation promoted by polyadenylation. Poster, S19-30. 6th International Congress of Plant Molecular Biology, Québec, Canada.

6.4 Acknowledgements

In the first place I want to thank PD Dr. Stefan Binder for giving me the opportunity to work on such a highly interesting topic. Stefan, thank you for your extraordinary support throughout the whole project, for many constructive discussions, and for being all ears whenever it was necessary. You guided me to work independently and to always view research results from several perspectives.

I want to thank Prof. Dr. Axel Brennicke for the outstanding working conditions in the department. Axel, I would like to thank you not only for your involvement in social matters and help with manuscripts, but also for your generous provision with cakes and pretzels :-).

Many thanks to Prof. Dr. Spindler, who agreed to engage with the topic and to revise this thesis.

I am very thankful to all current and former members of the „SB-Lab“ for creating a very funny and friendly surrounding, and for the aid in technical matters: Alex, Conny, Joachim, Manuela, Matthias, Nicole, Sabine, Saskia, and Stephie. Micha and Klaus, please accept my very thanks for your excellent suggestions, all the discussions and especially the hours beyond professional work. I would like to thank Bärbel and Ulrike for their extraordinary assistance.

A big thankyou to all members of the department of Molecular Botany for the informal atmosphere, the fruitful discussions, the technical help, all the every day services and the nice barbecue parties. Special thanks to Steffen and Anita for the standby with protein matters.

Sabine, I am very indebted to you for so many sunny moments as my bench neighbour in the lab and for being an island during the past year.

I am very much obliged to my family for enabling my studies and their financial as well as personal help throughout all the past years. I would like to dedicate this work to you.

I am very grateful to my friends for generating a perfect surrounding beyond the studies.

This work was financed by a scholarship of the Studienstiftung des deutschen Volkes.

6.5 Deutschsprachige Zusammenfassung

In der vorliegenden Arbeit wurden mehrere Gesichtspunkte der mRNA Prozessierung in Pflanzenmitochondrien untersucht. Dies geschah nicht nur auf Ebene der mRNA und ihrer Modifikationen, sondern auch durch die Charakterisierung eines Proteins, das möglicherweise an der 3' mRNA Prozessierung in Mitochondrien von höheren Pflanzen beteiligt ist.

Die identifizierte RNA Helikase aus *Arabidopsis thaliana* zeigt eine große Ähnlichkeit zu dem SUV3 Protein aus Hefe. Es konnte gezeigt werden, dass die im Zellkern kodierte DExH Box RNA Helikase (AtSUV3) in den Mitochondrien von *Arabidopsis thaliana* lokalisiert ist. Die AtSUV3 mRNA ist mit einer Länge von 2500 Nt aus 16 Exons zusammengesetzt und kodiert für ein Protein mit einer molekularen Masse von 63,6 kDa. Das Protein besitzt charakteristische Motive von DExH Box RNA Helikasen und zeigt eine schwache ATP spaltende Aktivität, die durch mitochondriale RNA stimulierbar ist, was ein deutlicher Hinweis auf eine Funktion als RNA Helikase ist.

Um den Einfluss posttranskriptionaler Modifikationen auf die 3' Prozessierung und die mRNA Stabilität in pflanzlichen Mitochondrien zu bestimmen, wurden das *atp9* Gen der Erbse und das *atp1* Gen aus *Oenothera* auf 3' nicht kodierte Nukleotide untersucht. Die gefundenen Polyadenylierungen befinden sich fast ausnahmslos an zuvor identifizierten 3' Enden. Funktionelle Analysen polyadenylierter Substrate der beiden Gene ergaben ein rasches Entfernen der homopolymeren Anhänge. Die in diesen RNAs vorhandenen Stem-Loop Strukturen verhindern den weiteren Abbau und heben fast vollständig die beschleunigende Wirkung der Polyadenylierung auf. Ganz im Gegensatz dazu wurden polyadenylierte Transkripte ohne eine stabilisierende Stem-Loop Struktur weit schneller abgebaut als Substrate ohne Poly(A)-Anhänge. Bei einer *in vivo* Untersuchung von 3' Enden in einem alternativen 3' RACE Ansatz wurden nicht kodierte Anhänge mit einer Länge von maximal 3 Nukleotiden identifiziert, die sich hauptsächlich aus Adenosinen zusammensetzten, die vorwiegend mit Cytidinen kombiniert sind. Weitere funktionelle Analysen von Substraten mit verschieden langen homopolymeren Anhängen zeigten, dass 3 Adenosine nicht für eine Stimulierung des Abbaus ausreichen, wohingegen 10 oder mehr Adenosine den Abbau deutlich beschleunigen. Generell scheint eine Polyadenylierung die Abbaurate von mRNA Molekülen zu beschleunigen, dem

jedoch die stabilisierende Wirkung einer Stem-Loop Struktur entgegenwirkt. Aus diesen antagonistischen Effekten resultiert eine effiziente Bildung von 3' prozessierten und stabilen Transkripten.

Bei der weiteren Charakterisierung von reifen Transkriptenden *in vivo* wurden für das *atp9* Gen aus der Erbse und das *atp1* Gen aus *Oenothera* Moleküle identifiziert, bei denen das 5' Ende kovalent mit dem 3' Ende verknüpft ist. Die Bildung solcher RNA Moleküle ist unabhängig von der Addition nicht kodierter Nukleotide, scheint aber von der Verfügbarkeit eines prozessierten 5' Endes abzuhängen. Ein solches 5' Ende zeichnet sich im Gegensatz zu einem primären transkriptionalen Ende durch eine Monophosphatgruppe aus, die für einer Ligationsreaktion Voraussetzung ist.

Um einen generelleren Einblick in die Bildung von Transkriptenden in Pflanzenmitochondrien zu erhalten, wurde das bereits zuvor untersuchte *cox2* Gen aus der Erbse exemplarisch mit Hilfe eines neuen methodischen Ansatzes untersucht. Dieser basiert auf einer Zirkularisierung der RNA und einer anschließenden RT-PCR über die Verknüpfungsstelle. Hierbei wurden die bereits bekannten Ergebnisse bezüglich der *cox2* Transkriptenden widerlegt. Die Verlässlichkeit dieser Methode wurde durch zusätzliche unabhängige Experimente bestätigt.

Erklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit selbständig angefertigt habe und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt sowie wörtlich oder inhaltlich übernommene Stellen als solche kenntlich gemacht habe.

Ulm, im Februar 2002

Josef Kuhn