

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,*}

^aDepartment of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK

^bAllgemeine 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

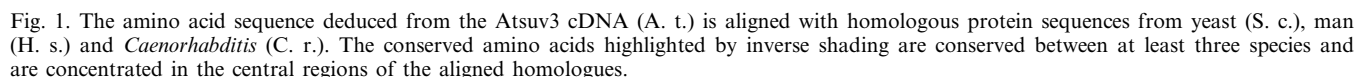
2.1. Identification of *suv3* homologue in *Arabidopsis thaliana* (*Atsuv3*)

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.



A fragment of the Atsvu3 cDNA encoding amino acids 11–387 was cloned into the pQE32 plasmid (Qiagen) to generate the plasmid pQE32-Atsvu3. A culture of XL1-Blue cells transformed with the pQE32-Atsvu3 plasmid was grown, induced and harvested according to the manual. Bacteria were disrupted by sonication in 20 mM Tris-

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 μ l of 20 mM Tris-HCl pH7.4, 50 mM NaCl, 3 mM MgCl₂, 1 mM DTT. Ten units of DNase I, 10 μ g of RNase A, 500 ng of yeast transfer RNA or 500 ng of total mitochondrial RNA were added to the reaction. Aliquots of 1 μ l 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]. This 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

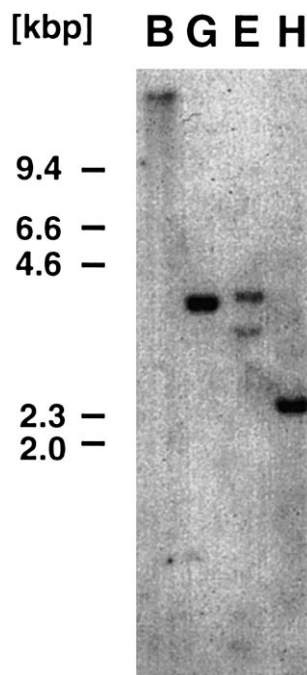


Fig. 2. The *Atsuv3* gene is a unique sequence in the *A. thaliana* nuclear genome. Restriction fragments identified with the genomic *Atsuv3* probe are with *Bam*HI (B) larger than 28 kb, with *Bgl*II (G) 4.114 kb, with *Eco*RI (E) 3.556, 4.251 and 0.734 kb and with *Hin*dIII (H) 2.566 kb respectively.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

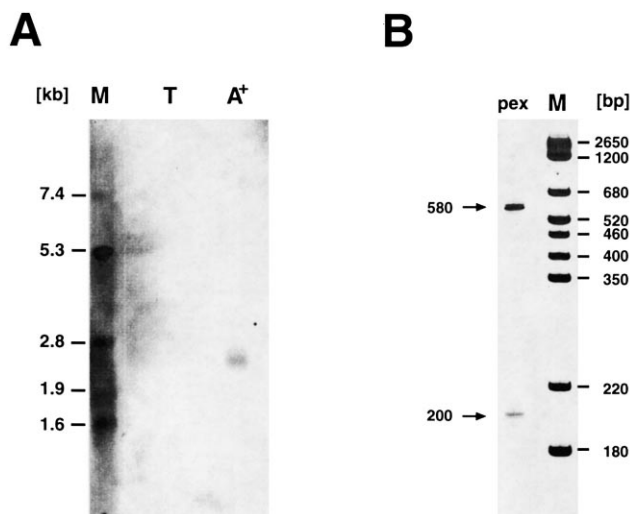


Fig. 3. The *Atsuv3* gene is transcribed at a very low level in *A. thaliana*. A: Northern blot analysis using 10 μ g of total RNA (lane T) and 5.7 μ g of poly(A)⁺ RNA (lane A⁺). Sizes of RNA marker molecules are given in kb. B: Primer extension analysis of 3 μ g poly(A)⁺ RNA using a primer located 40 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.

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].

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

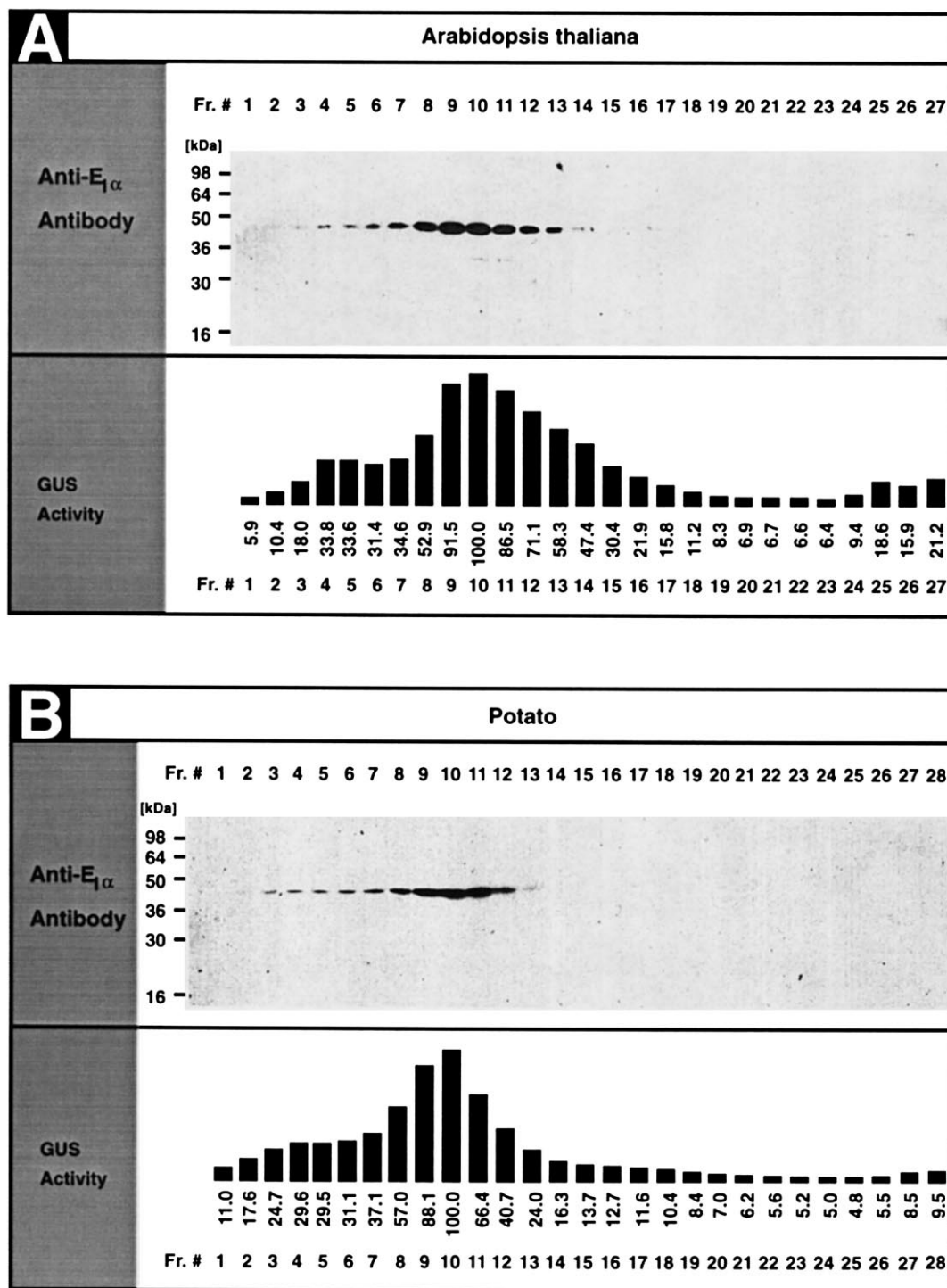


Fig. 4. The *Atsuv3* gene product is a mitochondrial protein. The intracellular localisation of the *Atsuv3*-encoded protein was analysed *in vivo* in transgenic *A. thaliana* (A) and potato (B) plants, which had been transformed with the first 387 codons of the *Atsuv3* open reading frame fused upstream 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.

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

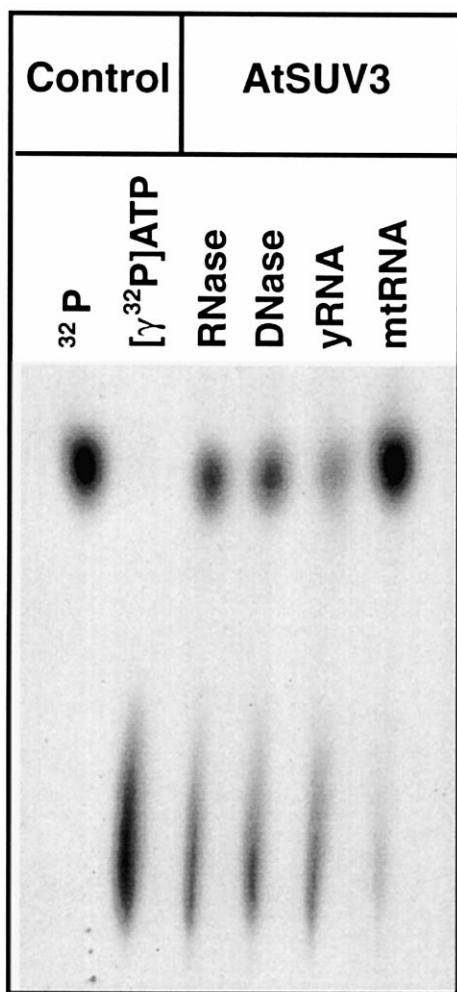


Fig. 5. ATP hydrolysis activity of the overexpressed and purified *AtSUV3* protein (amino acids 11–387) is stimulated by mitochondrial RNA. Hydrolysis of [γ -³²P]ATP was monitored by thin-layer chromatography, which separates phosphate (lane ³²P) from the slower migrating ATP (lane [γ -³²P]ATP). Addition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

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 Grisssem, 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 Grisssem, W. (1996) *EMBO J.* 15, 1132–1141.
- [4] Kudla, J., Hayes, R. and Grisssem, 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 Stepień, 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 Stepień, 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] Stepień, 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., Stepień, 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.