

# A chloroplastic RNA-binding protein is a new member of the PPR family

Sylvie Lahmy<sup>a</sup>, Fredy Barnèche<sup>a</sup>, Jean Derancourt<sup>b</sup>, Witold Filipowicz<sup>c</sup>, Michel Delseny<sup>a</sup>,  
Manuel Echeverria<sup>a,\*</sup>

<sup>a</sup>Laboratoire 'Génome et Développement des Plantes', Université de Perpignan, UMR CNRS 5096 Avenue de Villeneuve, 66860 Perpignan Cedex, France

<sup>b</sup>Centre de Recherches de Biochimie Macromoléculaire UPR1086 route de Mende, 34293 Montpellier Cedex 5, France

<sup>c</sup>Friedrich Miescher-Institut, P.O. Box 2543, CH-4002 Basel, Switzerland

Received 26 June 2000; revised 3 August 2000; accepted 3 August 2000

Edited by Gunnar von Heijne

**Abstract** P67, a new protein binding to a specific RNA probe, was purified from radish seedlings [Echeverria, M. and Lahmy, S. (1995) *Nucleic Acids Res.* **23**, 4963–4970]. Amino acid sequence information obtained from P67 microsequencing allowed the isolation of genes encoding P67 in radish and *Arabidopsis thaliana*. Immunolocalisation experiments in transfected protoplasts demonstrated that this protein is addressed to the chloroplast. The RNA-binding activity of recombinant P67 was found to be similar to that of the native protein. A significant similarity with the maize protein CRP1 [Fisk, D.G., Walker, M.B. and Barkan, A. (1999) *EMBO J.* **18**, 2621–2630] suggests that P67 belongs to the PPR family and could be involved in chloroplast RNA processing. © 2000 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

**Key words:** RNA-binding activity; PPR motif; Plant; Chloroplast; HA-tagging

## 1. Introduction

Regulation of RNA processing and translation play a major role in the selective control of both nuclear and organellar gene expression in higher plants. This has led to a growing interest in the identification of RNA–protein interactions that control these events. Many genetic and biochemical studies reveal the complexity and diversity of RNA–protein interactions. Specific interactions are dependent on different RNA structural elements. Their complexity is further increased by the presence within a single protein of different domains that cooperate to achieve specific RNA–protein interactions [3,4]. A general observation is that RNA-binding proteins often belong to large ribonucleoprotein complexes implicated in different RNA processing events.

In plants, many genes and cDNA encoding putative RNA-binding proteins have been identified by similarity to vertebrate or yeast proteins. Several classes of RNA-binding proteins have been described based on conserved RNA-binding motifs [5,6]. The majority of the plant RNA-binding proteins correspond to the RRM class [7,8] but putative DEAD-box RNA helicases have also been identified [9,10]. However, nat-

ural RNA substrates and functions are not known for most of these proteins [8].

We report here the cloning of a nuclear gene encoding a protein of 67 kDa previously isolated from radish and *Arabidopsis thaliana* using a pre-rRNA-binding assay and initially called NFC [1]. We demonstrated that P67 is an RNA-binding protein located in the chloroplast. It shows significant similarity to CRP1, a maize protein implicated in the processing and translation of specific chloroplast messenger RNAs [2,11]. CRP1 belongs to the so-called PPR family [12]. This suggests that P67 could be involved in the regulation of chloroplast RNA processing or translation.

## 2. Materials and methods

### 2.1. Plant material

Four day old radish seedlings (Vilmorin, National Rond rose à bout blanc) were used for protein purification. *A. thaliana* (Columbia ecotype) was used for other experiments.

### 2.2. Purification of P67/P60 and protein sequencing

Proteins P67 and P60 were extracted from radish seedlings and purified according to the published protocol [1]. The peak of P67/P60 RNA-binding activity was eluted by step on poly (U) Sepharose and represents the most purified P67/P60 fraction. The RNA-binding activity was followed by EMSA with an RNA probe. Ten µg of pure P67/P60 protein (obtained from 3.5 kg of homogenised tissue) were used for protein sequencing. After preparative electrophoresis, proteins were excised and prepared for in situ proteolysis according to Rosenfeld [13]. Resulting peptides were isolated by reverse phase HPLC on a C8 column (2×100 mm) eluted by an acetonitrile gradient in 0.1% trifluoroacetic. After a second purification on a C18 column, peptides were sequenced on a Procise sequencer (Perkin-Elmer, Foster City, CA, USA) using the manufacturer's pulsed liquid program.

### 2.3. Isolation of P67 genes

To clone the radish gene, degenerate primers derived from the P67 peptide sequences were used for PCR amplification on radish genomic DNA. Sequencing of the product allowed us to obtain the complete genomic sequence encoding P67 in radish by a PCR walking strategy [14]. Nested primers at the 5' and 3' ends of the fragment were used in combination with nested primers of the adapters for PCR amplification. The contig sequences were analysed and assembled using Sequencher (Genes Codes Corporation, Ann Arbor, MI, USA). Based on the sequence of the 5' and 3' regions flanking the radish P67 ORF, the full genomic sequence was amplified by PCR and cloned into the PCRscript plasmid (Stratagene). All plasmids were sequenced using the Prism Ready Reaction DyeDeoxy Terminator Cycle Sequencing kit and an ABI 373A automated DNA-sequencing apparatus (Applied Biosystems).

### 2.4. RT-PCR

Total RNA was isolated according to Kay et al. [15]. It was DNase-treated (1 h, 37°C, RQ1-DNase, Promega) to avoid genomic

\*Corresponding author. Fax: (33)-468668499.  
E-mail: manuel@univ-perp.fr

**Abbreviations:** PPR motif, pentatricopeptide motif; RRM, RNA recognition motif; rRNA, ribosomal RNA; EMSA, electrophoretic mobility shift assay; aa, amino acid; bZIP, basic leucine zipper

DNA contamination. 2.5 µg of total *Arabidopsis* RNA from different tissues were reverse-transcribed using RT-PCR kit (Stratagene). Aliquots of RNAs without reverse transcriptase were used as a control. One µl was used as a template in the subsequent PCR reaction. P67-5':5'-TCCAGAAACCTTCTCCAAGC-3' and P67-3':5'-CCAATTGGGAGTAAAGCCATTGG-3' were used for the amplification of P67. Actin mRNA (accession number U41998) was simultaneously amplified as a control.

## 2.5. Expression in *Escherichia coli*

The full-length P67 cDNA was amplified by PCR and sub-cloned into *NdeI/BamHI* sites of the *E. coli* expression plasmid pET-16b vector (Novagen) in frame with the N-terminal His-tag. BL21 *E. coli* containing the above plasmid were grown in LB medium at 37°C until OD<sub>600</sub> reached 0.5. Protein expression was induced by adding 1 mM IPTG and incubating 3 h at 30°C. After checking the protein induction on SDS-PAGE, cells were pelleted, resuspended in binding buffer (Novagen) and lysed by rounds of freezing and sonication.

## 2.6. EMSA analysis

Gel shift assays were performed as previously described [1] except that 0.1 µg yeast tRNA was added as competitor. A similar procedure was used with lysates of induced and non-induced bacteria except that 2 µg dI/dC and 2 µg yeast tRNA were added as competitors.

## 2.7. Plasmid constructs

The coding sequence of the *Rsl-P67* gene was cloned into the *NdeI/SalI* or the *BamHI/NdeI* sites of pHATO to obtain plasmids pHAT-P67 and pP67-HAT, expressing N- and C-HA-P67 respectively (HA-P67, P67-HA). The pHATO derived from pDEDH [16] contains a duplicated CaMV 35S promoter and a poly (A) signal to direct expression of a test protein tagged with an HA epitope in plant cells. HA corresponds to the influenza hemagglutinin non-peptide YPYDVPDYA.

## 2.8. Transfection of *Nicotiana plumbaginifolia* protoplasts and detection of HA-P67

Mesophyll protoplasts ( $3 \times 10^5$ ) of *N. plumbaginifolia* were transfected with pP67-HAT by the polyethylene glycol method [17] using 20 µg of plasmid per transfection. 24 h after transfection, immunolocalisation was carried out as described [18]. Immunodetection of P67-HA was performed with a rat monoclonal anti-HA (clone 3F10, Boehringer Mannheim). The anti-HA was revealed using fluorescein isothiocyanate-conjugated goat anti-rat IgG (Jackson ImmunoResearch Laboratories). No background was observed in plant cells when the primary antibody was not included (not shown). Slides were viewed with a fluorescence microscope (Zeiss Axiophot).

For Western blot analysis  $10^6$  *N. plumbaginifolia* protoplasts were transfected with plasmid pHAT-P67 or pP67-HAT. The pelleted protoplasts were resuspended in 100 µl of loading buffer. Ten µl was run on a 10% SDS-PAGE and proteins were transferred to a PVDF membrane (Amersham). The membrane was then blotted with a 1:100 dilution of mouse monoclonal anti-HA (clone 12CA5, Boehringer Mannheim). The immunoreactive proteins were detected using the ECL system from Amersham.

## 3. Results

### 3.1. Cloning of the genes encoding P67 in radish and *Arabidopsis*

We previously identified a 67 kDa protein and its proteolysed form, P60, from radish seedlings through pre-rRNA-binding activity [1]. After performing a large scale purification, five peptides from P67 (peptides 1–5 in Fig. 1) and two peptides from P60 were sequenced. The two peptides sequenced from P60 were identical to peptides 3 and 4 of P67, confirming that P60 is a modified form of P67 [1]. Primers derived from peptides 2 and 4 specifically amplified a 366 bp DNA fragment that also contained the coding sequence for peptide 3. By a PCR walking strategy [14] a genomic sequence of 2464 bp with an uninterrupted ORF of 2100 bp was ob-

tained. The conceptual translation product contains all the microsequenced peptides of P67. PCR amplification using radish genomic DNA as a template and specific primers flanking the ORF resulted in isolation of two genomic fragments that correspond to two different genes named *Rsl-P67* and *Rsl2-P67*. They encode nearly identical proteins (Fig. 1). Using the *Rsl-P67* sequence to screen the *Arabidopsis* databases, we identified an annotated gene on chromosome 4 (accession number Z97341, [19]), whose putative protein product aligned only with the C-terminal region of Rs-P67. Based on this sequence information, a 3122 bp genomic fragment from *Arabidopsis* was amplified and fully sequenced. An error in the first deposited sequence had introduced an insertion preventing correct annotation. An uninterrupted ORF of 2108 bp encoding a protein having 87% identity to the *Rsl-P67* was identified (Fig. 1). Indeed, two genes in radish and one in *Arabidopsis* were detected by Southern experiments (not

## A

|          |     |                                              |
|----------|-----|----------------------------------------------|
| Rsl-P67  | 1   | MSFHLIYSS- PSSLHDPYPLCNLLSVHHKSTPRSFVSSYNPN  |
| Rsl2-P67 | 1   | MSFHLIYSS- PSSLHDPYPLCNLLSVHHKSTPRSFVSSYNPN  |
| At-P67   | 1   | MSFHLIYSS- PSSLHDPYPLCNLLSVHHKSTPRSFVSSYNPN  |
| Rsl-P67  | 42  | SPPFHSRTLLQTSVLSLOEPLPQETQTEKPELIDANPPAS- GS |
| Rsl2-P67 | 42  | SPPFHSRTLLQTSVLSLOEPLPQETQTEKPELIDANPPAS- GS |
| At-P67   | 42  | SPPFHSRTLLQTSVLSLOEPLPQETQTEKPELIDANPPAS- GS |
| Rsl-P67  | 83  | KRYAWVNPSPRASOLRRKSYDSRYSSLVKLAESLDSCLPNE    |
| Rsl2-P67 | 83  | KRYAWVNPSPRASOLRRKSYDSRYSSLVKLAESLDSCLPNE    |
| At-P67   | 83  | KRYAWVNPSPRASOLRRKSYDSRYSSLVKLAESLDSCLPNE    |
| Rsl-P67  | 125 | ADVSDVIAKFGSKLFEODAVVTLNNMNPETAPLVNLLLET     |
| Rsl2-P67 | 125 | ADVSDVIAKFGSKLFEODAVVTLNNMNPETAPLVNLLLET     |
| At-P67   | 125 | ADVSDVIAKFGSKLFEODAVVTLNNMNPETAPLVNLLLET     |
| Rsl-P67  | 167 | LKPTREVILYNVTMKVFRKSKDLEKSEKLFDEMLQRGVKKPD   |
| Rsl2-P67 | 167 | LKPTREVILYNVTMKVFRKSKDLEKSEKLFDEMLQRGVKKPD   |
| At-P67   | 167 | LKPTREVILYNVTMKVFRKSKDLEKSEKLFDEMLQRGVKKPD   |
| Rsl-P67  | 209 | ATFTTLISCAROCGLPRAVWEFEKMPSEFGLPNDVTLAAMI    |
| Rsl2-P67 | 209 | ATFTTLISCAROCGLPRAVWEFEKMPSEFGLPNDVTLAAMI    |
| At-P67   | 209 | ATFTTLISCAROCGLPRAVWEFEKMPSEFGLPNDVTLAAMI    |
| Rsl-P67  | 251 | DAYGRAGNVEMALSLYDRARTEKWRIDPVTFTSLIRIYGYAG   |
| Rsl2-P67 | 251 | DAYGRAGNVEMALSLYDRARTEKWRIDPVTFTSLIRIYGYAG   |
| At-P67   | 251 | DAYGRAGNVEMALSLYDRARTEKWRIDPVTFTSLIRIYGYAG   |
| Rsl-P67  | 293 | NYDGCNLIYEEEMKSLGVKPNLVINRLDSMGKAKRPWOATM    |
| Rsl2-P67 | 293 | NYDGCNLIYEEEMKSLGVKPNLVINRLDSMGKAKRPWOATM    |
| At-P67   | 293 | NYDGCNLIYEEEMKSLGVKPNLVINRLDSMGKAKRPWOATM    |
| Rsl-P67  | 335 | LHKDLISNGFEPNWSYAAALRAYGRARYGEDALVIRYOMKG    |
| Rsl2-P67 | 335 | LHKDLISNGFEPNWSYAAALRAYGRARYGEDALVIRYOMKG    |
| At-P67   | 335 | LHKDLISNGFEPNWSYAAALRAYGRARYGEDALVIRYOMKG    |
| Rsl-P67  | 377 | KGLELTIVLYNTLLSMCADIGYVDEAFEIFODMKSGTCEPD    |
| Rsl2-P67 | 377 | KGLELTIVLYNTLLSMCADIGYVDEAFEIFODMKSGTCEPD    |
| At-P67   | 377 | KGLELTIVLYNTLLSMCADIGYVDEAFEIFODMKSGTCEPD    |
| Rsl-P67  | 419 | SWTFSSLLITVYSCGRVSEAEALREMRAGEFPTFLVLTSL     |
| Rsl2-P67 | 419 | SWTFSSLLITVYSCGRVSEAEALREMRAGEFPTFLVLTSL     |
| At-P67   | 419 | SWTFSSLLITVYSCGRVSEAEALREMRAGEFPTFLVLTSL     |
| Rsl-P67  | 461 | IOCYGKAKOVDDVVRTFEQVLELGEIPEDDRCGCLLNVMTOT   |
| Rsl2-P67 | 461 | IOCYGKAKOVDDVVRTFEQVLELGEIPEDDRCGCLLNVMTOT   |
| At-P67   | 461 | IOCYGKAKOVDDVVRTFEQVLELGEIPEDDRCGCLLNVMTOT   |
| Rsl-P67  | 503 | PIEEIGKLIIGVEKAKPKLGRVYKMLVEEENCEEGLVKKKAS   |
| Rsl2-P67 | 503 | PIEEIGKLIIGVEKAKPKLGRVYKMLVEEENCEEGLVKKKAS   |
| At-P67   | 503 | PIEEIGKLIIGVEKAKPKLGRVYKMLVEEENCEEGLVKKKAS   |
| Rsl-P67  | 545 | ELIDSIGSDVNKAYLNCILDLVNLNLEKACEILQGLGYD      |
| Rsl2-P67 | 545 | ELIDSIGSDVNKAYLNCILDLVNLNLEKACEILQGLGYD      |
| At-P67   | 545 | ELIDSIGSDVNKAYLNCILDLVNLNLEKACEILQGLGYD      |
| Rsl-P67  | 587 | IYSGLOSKSATOWSLHLKSLSLGAALTALHYVWMDLSEAAIT   |
| Rsl2-P67 | 587 | IYSGLOSKSATOWSLHLKSLSLGAALTALHYVWMDLSEAAIT   |
| At-P67   | 587 | IYSGLOSKSATOWSLHLKSLSLGAALTALHYVWMDLSEAAIT   |
| Rsl-P67  | 629 | SGEEFPPLLGINTGHHGKHKYSKGLAAVFESHKLKELNAPFHE  |
| Rsl2-P67 | 629 | SGEEFPPLLGINTGHHGKHKYSKGLAAVFESHKLKELNAPFHE  |
| At-P67   | 629 | SGEEFPPLLGINTGHHGKHKYSKGLAAVFESHKLKELNAPFHE  |
| Rsl-P67  | 671 | APDKVGWFLTTSVAAKTWLESRRSSAEVSA               |
| Rsl2-P67 | 671 | APDKVGWFLTTSVAAKTWLESRRSSAEVSA               |
| At-P67   | 671 | APDKVGWFLTTSVAAKTWLESRRSSAEVSA               |

## B

245 - TLAANIDAGRAGNVEMALSPDRARTEKWRIDP 249  
 349 - TLAANIDAGRAGNVEMALSPDRARTEKWRIDP 383  
 PPR consensus: TNAANIDAGRAGNVEMALSPDRARTEKWRIDP

Fig. 1. Alignment of the amino acid sequences of P67 in radish and *A. thaliana*. A: Sequences were deduced from the ORF encoded by the radish *Rsl-P67*, *Rsl2-P67* and the *Arabidopsis* *At-P67* genes (accession numbers A5243544, A5243546, A5243545 respectively). Boxes indicate identical residues and shading similar ones. The five sequenced peptides are indicated in bold type. The putative chloroplast transit peptide is indicated by dashes. Two PPR motifs are overlaid by arrows. The four leucines constituting the predicted leucine zipper are shown by circles. B: Comparison of P67 motifs to PPR consensus [12].

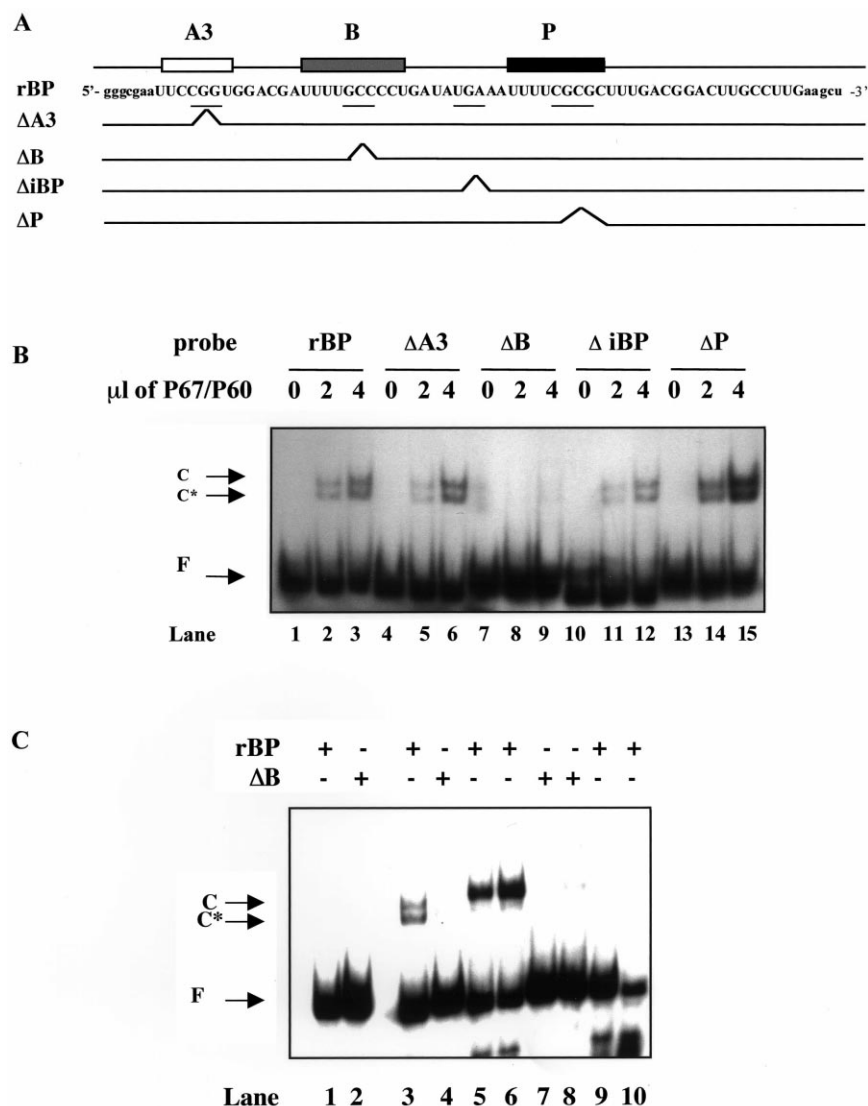


Fig. 2. Radish P67 and P67r bind specifically to rBP fragment. A: The sequences of the rBP and the mutated probes are shown. The upper case letters indicate nucleotide sequence from the radish pre-rRNA which contains three similar motifs A<sup>3</sup>, B and the cleavage site P. Lower case letters show nucleotides derived from plasmid polylinker used for in vitro transcription. Deleted nucleotides are indicated by horizontal dashes. B: The binding of P67 to the RNA fragments detected by EMSA was performed with the indicated amount of pure P67/P60 fraction and 2 fmol of each indicated probe, in the presence of 2 μg of tRNA. C: The EMSA was carried out as in Section 2 with the rBP or ΔB fragments and the following extracts: no protein (lanes 1 and 2), 4 μl of radish P67/P60 fraction (lanes 3 and 4), 2 μl (lanes 5 and 7) and 4 μl (lanes 6 and 8) of bacterial lysates from induced cells transfected with pET-P67, 4 μl of uninduced cells transfected with pET-P67 (lane 9), 4 μl of bacterial lysate from induced cells transfected with pET vector (lane 10). Bands migrating faster than F (free probe) correspond to degraded probe. F, free probe; C and C\*, RNA/P67 and RNA/P60 complexes respectively.

shown). We do not know whether one or both radish genes are expressed.

The ORFs of *RsI-P67* and the *At-P67* genes encode proteins of 700 and 702 aa respectively (78 kDa, pI of 5.7). Several sequence motifs were identified in radish and *Arabidopsis* P67. The N-terminal 55 aa region exhibits features of a chloroplast transit peptide [20]. Two similar stretches of 35 amino acids were found in the region 244–384. These repeats show strong similarity to the PPR motifs described recently in plant organellar proteins [12]. A four heptad leucine zipper is also predicted at the carboxy-terminus of the protein (Fig. 1) that could mediate protein–protein interactions [21]. Unlike transcription factors of the bZIP family [22], P67 does not contain a basic domain next to the leucine zipper. No known RNA-binding motif could be detected.

Alignment of P67 with databases revealed no significant similarity to eukaryotic or prokaryotic proteins [23] except with the chloroplast maize protein CRP1 [2,11] which has 18% identity and 53% similarity to P67 distributed throughout the sequence (not shown).

### 3.2. P67 recombinant protein has an RNA-binding activity

To obtain additional evidence that identified genes indeed encode the P67 protein, we tested the RNA-binding activity of the recombinant P67 (P67r) in the bacterial lysates by EMSA. We have shown that P67 and its related form P60 specifically bind to an RNA fragment (rBP) encompassing the first cleavage site of the radish pre-rRNA [1]. Predicted RNA secondary structures suggest that the rBP fragment is highly structured and could fold into two stems formed by B and P motifs

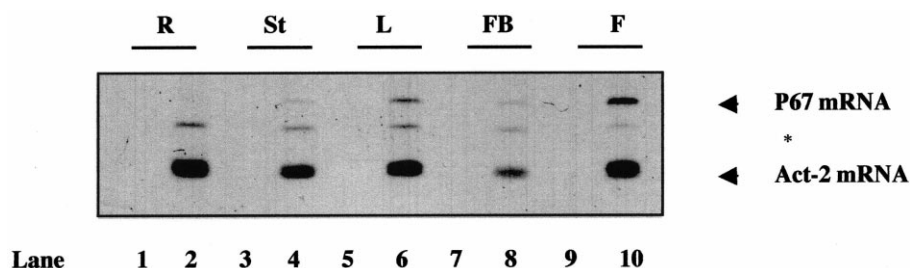


Fig. 3. Expression of P67 mRNA. *A. thaliana* roots (R), stems (St), leaves (L), flower buds (FB) and flowers (F). Even lanes correspond to expression of P67 and actin-2. Uneven lanes: reverse transcriptase was omitted in the first strand cDNA synthesis. \* corresponds to an artifactual amplification product.

separated by a large loop (Mfold program [24], not shown). Deletions that alter the predicted structure of the rBP fragment to various extents were prepared (Fig. 2A). rBP,  $\Delta A3$ ,  $\Delta iBP$ ,  $\Delta P$  could fold into similar structures but  $\Delta B$  was completely different (not shown). The binding activity of P67 to  $\Delta A3$ ,  $\Delta iBP$  and rBP fragments are identical (Fig. 2B, lanes 2, 3 and lanes 5, 6, 11, 12). Deletion in the P motif slightly increased the binding of P67 to the probe (Fig. 2B, lanes 2, 3 and lanes 14, 15). In contrast  $\Delta B$  completely prevents binding (Fig. 2B, lanes 2, 3 and lanes 8, 9). These data suggest that the binding of P67 to rBP has specific structural requirements.

Since the most significant difference in binding was detected with rBP and  $\Delta B$  fragments, the binding of P67r to these probes was analysed and compared to that of purified P67/P60. A specific binding to rBP fragments is detected with extracts from induced cells (Fig. 2C, lanes 5, 6), producing a shifted band having similar mobility to the complex generated by the purified P67/P60 fraction (Fig. 2C, lane 3). Such binding is prevented by using the  $\Delta B$  probe (Fig. 2C, lane 4 and lanes 7, 8). The C\* complex formed with the plant extracts is not observed with the recombinant protein fraction (Fig. 2C, lane 3 and lanes 5, 6). This suggests that P60 is either produced during the extraction procedure by proteolysis or is a natural modification of P67 occurring in the plant cell. No binding was detected with extracts either from uninduced cells or from cells transformed with the empty vector (Fig. 2C, lanes 9 and 10).

These results confirm that the *Rsl-P67* gene encodes the RNA-binding protein P67 purified from plant cells.

### 3.3. Expression of P67 mRNA in *A. thaliana*

RT-PCR was used to detect P67 mRNA in different *Arabidopsis* tissues. It revealed that P67 mRNA is expressed in leaves and flowers (Fig. 3, lanes 6 and 10) and at lower levels in stems and flower buds (Fig. 3, lanes 4 and 8); no expression was found in roots. Actin mRNA was simultaneously analysed as a control.

### 3.4. Subcellular localisation of P67

Sequencing of the gene indicated that the N-terminus of P67 has typical features of chloroplast transit peptides, thus raising the question of the subcellular localisation of P67. Epitope HA-tagged proteins were engineered and expressed in protoplasts. Western blot analysis with a mouse anti-HA indicated that a 67 kDa protein is specifically detected in the extracts prepared from protoplasts expressing the P67-HA (Fig. 4B, lane 2) but not HA-P67 (Fig. 4B, lane 3). Considering that both constructs only differ in the position of the

HA-tag, the failure to detect HA-P67 probably reflects the removal of the N-terminal HA-tag in vivo. We directly visualised the subcellular localisation of P67-HA by immunofluorescence. In the transfected protoplasts, expression of P67-HA is associated with chloroplasts and is clearly above the background of a cell not expressing P67-HA (Fig. 4C, parts B and E). In all cases the P67-HA green fluorescence is specifically associated with the chloroplasts labelled by the red autofluorescence of chlorophyll (Fig. 4C, parts A and D). Transient expression of P67-GFP in *N. plumbaginifolia* protoplasts has also shown that P67 is localised in chloroplasts (data not shown).

## 4. Discussion

Similarity searches in databases using radish and *Arabidopsis* P67 sequences suggest that P67 is a previously undescribed plant RNA-binding protein. Though we first identified this protein in radish seedlings by its interaction with a pre-rRNA probe, analysis of the P67 gene indicated that a 78 kDa precursor is processed into a 67 kDa mature form and addressed to chloroplasts (Fig. 4). Its expression level is low and only observed in the aerial part of the plant, which is in agreement with its presence in photosynthetic tissues.

An intriguing feature of this chloroplast protein is its highly specific RNA-binding activity: in vitro it only recognises the rBP RNA and the recombinant protein has the same characteristics (Fig. 2). However, the presence of the protein in the chloroplast clearly indicates that the nuclear pre-rRNA is not the natural substrate of this protein. No sequence similar to rBP was found in the sequenced chloroplast genomes [25]. We suspect that the specific interaction of P67 with rBP is directed by secondary or tertiary structures of the RNA fragment, as reported for many RNA-binding proteins [6]. Significantly, it has been shown that in chloroplasts the differential stability and translation of some *cp*rRNA is controlled through *cis*-acting elements forming stem-loop structures in their untranslated 5' and 3' regions which are targets for specific RNA-binding proteins [26–28]. Thus it is not unlikely that some of these structures could be putative substrates for P67.

P67 contains no known conserved RNA-binding motif. A good candidate for the RNA-binding domain could be the PPR motif as recently suggested [12]. Some of them might be RNA-binding rather than protein-binding motifs. Another interesting motif is the predicted leucine-zipper motif in the carboxy-terminus of P67. In plants, no RNA-binding proteins with leucine-zipper motifs have been reported so far and usually bZIP factors are DNA-binding proteins [29,30]. RNA-

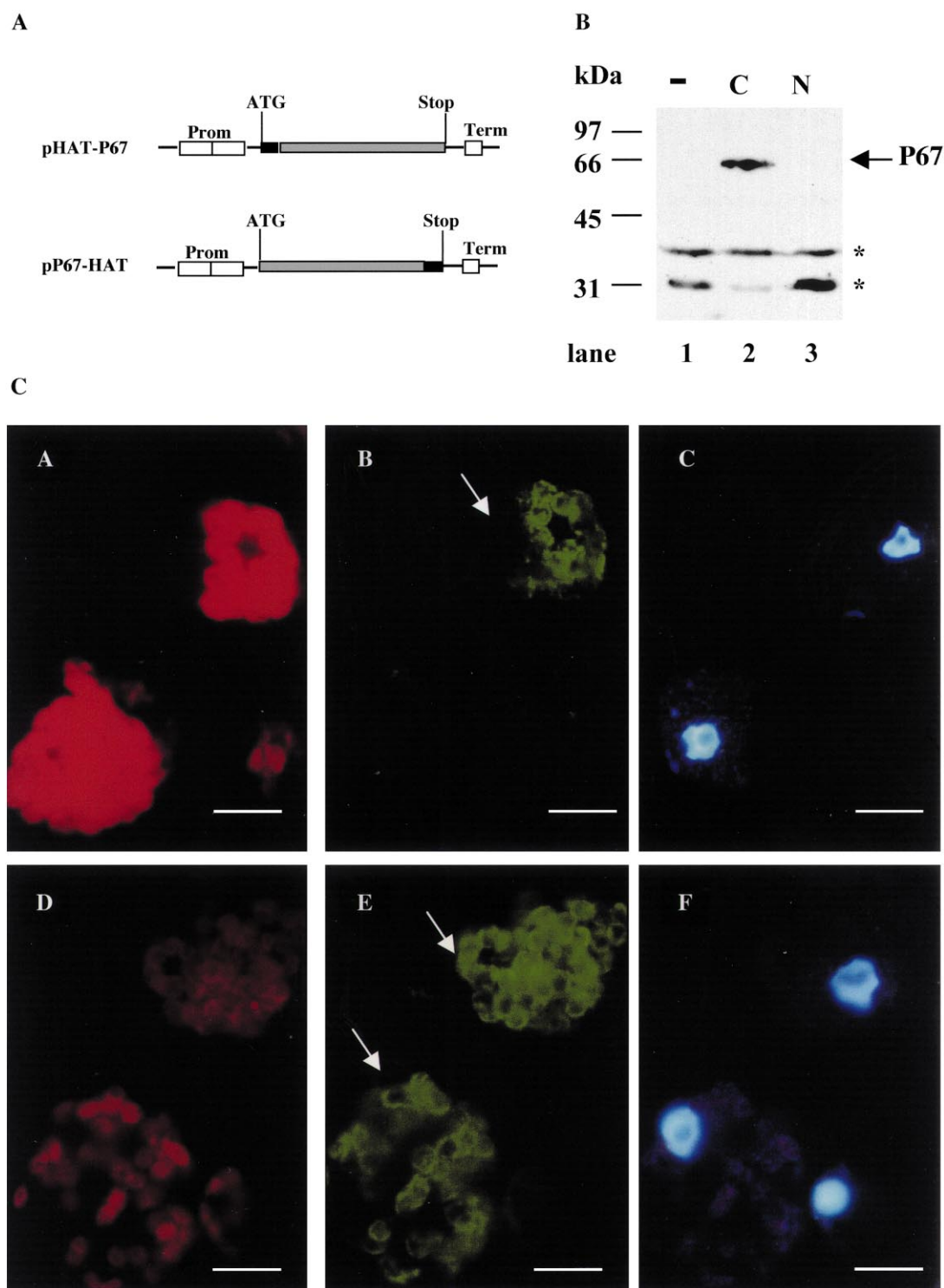


Fig. 4. Transient expression of HA-P67 in transfected *N. plumbaginifolia* protoplasts. A: Schematic diagram (not to scale) of constructs in pHATO vector used for transient expression in protoplasts. The pHAT-67 and pP67-HAT induced expression of N-tagged and C-tagged HA-P67 proteins respectively. Clear boxes (Prom and Term) indicate the duplicated CaMV 35S promoter and CaMV poly (A) signal respectively. Grey boxes indicate the coding sequence of *Rsl-P67*; the black box represents the HA epitope recognised by anti-HA antibodies. B: Expression of P67-HA (C, lane 2) or HA-P67 (N, lane 3) was detected with the mouse monoclonal anti-HA. A control was made with extracts from protoplasts transfected with pHATO (lane 1). \* 30 and 40 kDa non-specific bands of tobacco proteins cross-reacting with anti-HA. C: Chloroplast localisation of P67: Cells were immunostained with HA antibody and fluorescein goat anti-rat IgG as the secondary antibody (B and E, green fluorescence under blue light excitation) The red fluorescence corresponds to chlorophyll emission (A and D). DAPI staining nuclei allow to localise cells (C and F). Arrows indicated the transfected cells. The bar represents 20 μm.

binding proteins with leucine-zipper motifs have only been described in animals [31]. For instance in mouse the leucine zipper of TB-RBP protein directs the dimerisation which is essential for its RNA-binding activity [32,33]. Since homodimerisation of P67 is unlikely [1], the leucine zipper could be involved in heterodimer formation.

The best similarity score was found with the CRP1 protein which also belongs to the PPR protein group. CRP1 is involved in both processing of *petD* *cp*RNA and translation of *petA* and *petD* *cp*RNAs [11]. In plants and green algae post-transcriptional regulation of chloroplast gene expression plays a major role in the control of plastid biogenesis and development: differential transcript stability regulates *cp*RNA accumulation during plastid development, while translation is strictly controlled by light [34]. These events are mediated by RNA structures of the 5'-UTR and 3'-UTR of chloroplast transcripts [26,27,35].

Considering these data, we propose that P67 interacts with chloroplastic proteins to form a complex involved in the processing or the translation of *cp*RNAs.

**Acknowledgements:** We specially thank Drs R. Cooke and Y. Meyer for help in computer programs and sequence alignments. We are grateful to Michèle Laudé and Christel Berger for their technical help with the automatic sequencer. English was kindly corrected by Dr R. Cooke. We are also grateful to I. Petruska for protoplast transfection and to D. Wicczorek for assistance with immunolocalisation. This work was supported by CNRS grant UMR5096.

## References

- [1] Echeverria, M. and Lahmy, S. (1995) *Nucleic Acids Res.* 23, 4963–4970.
- [2] Fisk, D.G., Walker, M.B. and Barkan, A. (1999) *EMBO J.* 18, 2621–2630.
- [3] Serin, G., Joseph, G., Ghisolfi, L., Bauzan, M., Erard, M., Amalric, F. and Bouvet, P. (1997) *J. Biol. Chem.* 272, 13109–13116.
- [4] Chandler, S.D., Mayeda, A., Yeakley, J.M., Krainer, A.R. and Fu, X.D. (1997) *Proc. Natl. Acad. Sci. USA* 94, 3596–3601.
- [5] Burd, C.G. and Dreyfuss, G. (1994) *Science* 265, 615–620.
- [6] Draper, D.E. (1995) *Annu. Rev. Biochem.* 64, 593–620.
- [7] Hecht, V., Stiefel, V., Delseny, M. and Gallois, P. (1997) *Plant Mol. Biol.* 34, 1119–1124.
- [8] Alba, M.M. and Pages, M. (1998) *Trends Plant Sci.* 3, 15–21.
- [9] Aubourg, S., Kreis, M. and Lecharny, A. (1999) *Nucleic Acids Res.* 27, 628–636.
- [10] Lorkovic, Z.J., Herrmann, R.G. and Oelmüller, R. (1997) *Mol. Cell. Biol.* 17, 2257–2265.
- [11] Barkan, A., Walker, M., Nolasco, M. and Johnson, D. (1994) *EMBO J.* 13, 3170–3181.
- [12] Small, I.D. and Peteers, N. (2000) *Trends Biochem. Sci.* 25, 46–47.
- [13] Rosenfeld, J., Capdevielle, J., Guillemot, J.C. and Ferrara, P. (1992) *Anal. Biochem.* 203, 173–179.
- [14] Devic, M., Albert, S., Delseny, M. and Roscoe, T.J. (1997) *Plant Physiol. Biochem.* 35, 331–339.
- [15] Kay, R., Chan, A., Daly, M. and McPherson, C.W. (1987) *Science* 236, 1299–1302.
- [16] Goodall, G.J. and Filipowicz, W. (1989) *Cell* 58, 473–483.
- [17] Goodall, G.J., Wiebauer, K. and Filipowicz, W. (1990) *Methods Enzymol.* 181, 148–161.
- [18] Heinlein, M., Epel, B.L., Padgett, H.S. and Beachy, R.N. (1995) *Science* 270, 1983–1985.
- [19] Bevan, M., Bancroft, I., Bent, E., Love, K., Goodman, H., Dean, C., Bergkamp, R., Dirkse, W., Van Staveren, M., Stiekema, W., Drost, L., Ridley, P., Hudson, S.A., Patel, K., Murphy, G., Piffanelli, P., Wedler, H., Wedler, E., Wambutt, R., Weitzenegger, T., Pohl, T.M., Terry, N., Gielen, J., Villarroel, R. and Chaltzisz, N. et al. (1998) *Nature* 391, 485–488.
- [20] Cline, K. and Henry, R. (1996) *Annu. Rev. Cell Dev. Biol.* 12, 1–26.
- [21] Cohen, C. and Parry, D.A. (1994) *Science* 263, 488–489.
- [22] Hurst, H.C. (1994) in: *Protein Profile, Transcription Factors 1: bZIP Proteins*, Vol. 1, pp. 123–168, Academic Press, New York.
- [23] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) *J. Mol. Biol.* 215, 403–410.
- [24] Zucker, M., Mathews, D.H. and Turner, D.H. (1999) in: *RNA Biochemistry and Biotechnology* (Barciszewski, J. and Clark, B.F.C., Eds.), pp. 11–43, Kluwer Academic Publishers, Dordrecht.
- [25] Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E. and Tabata, S. (1999) *DNA Res.* 6, 283–290.
- [26] Hayes, R., Kudla, J., Schuster, G., Gabay, L., Maliga, P. and Gruissem, W. (1996) *EMBO J.* 15, 1132–1141.
- [27] Klaff, P., Mundt, S.M. and Steger, G. (1997) *RNA* 3, 1468–1479.
- [28] Yohn, C.B., Cohen, A., Danon, A. and Mayfield, S.P. (1998) *Proc. Natl. Acad. Sci. USA* 95, 2238–2243.
- [29] Oyama, T., Shimura, Y. and Okada, K. (1997) *Genes Dev.* 11, 2983–2995.
- [30] Fulgosi, H., Vener, A.V., Altschmied, L., Herrmann, R.G. and Andersson, B. (1998) *EMBO J.* 17, 1577–1587.
- [31] McAfee, J.G., Shahied-Milam, L. and Soltaninassab, S.R. (1996) *RNA* 2, 1139–1152.
- [32] Wu, X.Q., Gu, W., Meng, X. and Hecht, N.B. (1997) *Proc. Natl. Acad. Sci. USA* 94, 5640–5645.
- [33] Wu, X.Q., Xu, L. and Hecht, N.B. (1998) *Nucleic Acids Res.* 26, 1675–1680.
- [34] Gruissem, W. and Tonkyn, J.C. (1993) *Crit. Rev. Plant Sci.* 12, 19–55.
- [35] Danon, A. and Mayfield, S.P. (1994) *EMBO J.* 13, 2227–2235.