

A functional RNase P protein subunit of bacterial origin in some eukaryotes

Lien B. Lai (1, 2), Pilar Bernal-Bayard (3), Gireesha Mohannath (1, 2), Stella M. Lai (1, 2), Venkat Gopalan (1, 2) and Agustín Vioque (3)

(1)Department of Biochemistry, The Ohio State University, Columbus, OH 43210, USA

(2)Center for RNA Biology, The Ohio State University, Columbus, OH 43210, USA

(3)Instituto de Bioquímica Vegetal y Fotosíntesis, Universidad de Sevilla and CSIC, Américo Vespucio 49, 41092 Sevilla, Spain

Abstract

RNase P catalyzes 5'-maturation of tRNAs. While bacterial RNase P comprises an RNA catalyst and a protein cofactor, the eukaryotic (nuclear) variant contains an RNA and up to ten proteins, all unrelated to the bacterial protein. Unexpectedly, a nuclear-encoded bacterial RNase P protein (RPP) homolog is found in several prasinophyte algae including *Ostreococcus tauri*. We demonstrate that recombinant *O. tauri* RPP can functionally reconstitute with bacterial RNase P RNAs (RPRs) but not with *O. tauri* organellar RPRs, despite the latter's presumed bacterial origins. We also show that *O. tauri* PRORP, a homolog of Arabidopsis PRORP-1, displays tRNA 5'-processing activity in vitro. We discuss the implications of the striking diversity of RNase P in *O. tauri*, the smallest known free-living eukaryote.

Keywords

RNase P diversity – Prasinophyte – pre-tRNA processing

Introduction

Despite a shared primary function in 5'-end tRNA maturation, the make-up of the RNase P ribonucleoprotein (RNP) complex is evolutionarily divergent (Evans et al. 2006; Walker and Engelke 2006; Esakova and Krasilnikov 2010; Jarrous and Gopalan 2010; Lai et al. 2010; Liu and Altman 2010). While there is a single catalytic RNase P RNA (RPR) subunit that shares a common ancestry in all three domains of life, the RNase P protein (RPP) component is variable in number and identity. In Bacteria, RNase P is associated with a single protein subunit (~14 kDa) while up to five RPPs have been reported in Archaea. Eukarya has up to ten RPPs; five of them are homologous to archaeal RPPs, but none have homology to the bacterial RPP.

Organellar RNase P presents a more complicated scenario. Despite the presence of a bacterial RPR-like gene in some algal mitochondria and chloroplast genomes (Lai et al. 2010), no bacterial RPP homolog has been reported in eukaryotes. The organellar RPPs characterized in yeast and human mitochondria are distinct from any of the known RPPs (Morales et al. 1992; Holzmann et al. 2008). Although a pioneering study demonstrated that human mitochondrial (mt) RNase P is devoid of RNA and functions as a patchwork of three proteins (Holzmann et al. 2008), recent work suggests the co-existence of both a protein-only and an RNP-based RNase P (Wang et al. 2010). In another breakthrough, the organellar precursor tRNA (pre-tRNA) 5'-processing activity in Arabidopsis was shown to be associated with a single nuclear-encoded polypeptide (PRORP-1), which is homologous to one of the three proteins in the protein-only form of human mt RNase P (Gobert et al. 2010).

Given this extraordinary diversity in the subunit make-up of RNase P, we inquired into the RNase P variants that might be present in *Ostreococcus tauri*, a unicellular green alga. Our focus on *O. tauri* was inspired by three reasons: it is the smallest known free-living eukaryote; it is a member of the Prasinophyceae, an ancient sister clade to land plants; and its nuclear and organellar genome sequences are available. Moreover, since its nuclear genome is one of the smallest (12.5 Mb) and most compact (1.5 kb/gene) among eukaryotes, *O. tauri* presents a unique opportunity to inquire if there is any correlation between the number of RNase P variants and genome size. The recently sequenced nuclear genomes of *O. tauri* and four other prasinophytes (Derelle et al. 2006; Palenik et al. 2007; Worden et al. 2009) unexpectedly reveal the presence of a hypothetical protein that is homologous to the bacterial RPP. We demonstrate here that a recombinant version of this bacterial-like RPP from *O. tauri* assembles with bacterial RPRs to generate holoenzymes capable of tRNA 5'-processing, a finding which represents the first discovery of an active bacterial-like RNase P protein subunit in eukaryotes. Although we document evidence for the expression of *O. tauri* RPP in its native context, the biological partners and function of *O. tauri* RPP (or its prasinophyte homologs) remain to be identified. We also found that a recombinant *O. tauri* PRORP, a homolog of *Arabidopsis* PRORP-1, displays pre-tRNA cleavage in vitro. We discuss the implications of the diversity of RNase P isoforms in *O. tauri*.

Materials and methods

Culturing of *O. tauri*

O. tauri RCC745 cells were obtained from Roscoff Culture Collection and grown on K medium at 25°C (Keller et al. 1987).

Cloning, expression and purification of *O. tauri* RPP

The coding region of *O. tauri* RPP was amplified by PCR from genomic DNA using OtC5-2F (5'-GATCCATGGCGTCGGGGGGTGGG-3') and OtC5-2R (5'-GTCACGTGTTGGTCGGCGTCGGTCTGGC-3'). The NcoI and PmlI recognition sites (underlined) introduced by the primers facilitated cloning of the PCR product into pET-33b at the NcoI and filled-in EagI sites, respectively, to create a C-terminal His6-tag fusion. The resulting plasmid (pET-33b-OtRPP) was confirmed by sequencing before it was used to transform *Escherichia coli* BL21(DE3) Rosetta cells (Novagen). A single transformant was grown to Abs600 ~0.6 at 37°C in LB medium supplemented with 35 µg/mL chloramphenicol + 35 µg/mL kanamycin and induced with 0.5 mM IPTG for 20 h at room temperature (~23°C). For purification, a 125 mL culture pellet was resuspended in 12.5 mL Buffer A [20 mM sodium phosphate (pH 7.2), 4 M urea, 0.2 mM PMSF], sonicated, and centrifuged at 30,000g for 20 min at 4°C. The crude extract was passed through a 0.45 µm syringe filter and loaded onto a 1 mL HiTrap CM-Sepharose column (GE Healthcare). All subsequent solutions for purification contained 20 mM sodium phosphate, pH 7.2. Using an FPLC (Pharmacia) apparatus, a linear reverse-urea gradient (4–0 M urea) was employed to renature the proteins before their elution with a linear 0–1.5 M NaCl gradient. The peak fractions (~0.75 to 1.1 M NaCl) were pooled, diluted to 0.5 M NaCl, and loaded onto a 1 mL HiTrap Chelating column (GE Healthcare) equilibrated with NiSO₄. When a linear 0–0.5 M imidazole gradient was used, we observed co-elution of three proteins at ~0.35 M imidazole. To further purify the proteins individually, the peak imidazole fractions were pooled and subjected to a 12% (w/v) polyacrylamide-SDS gel electrophoresis. Each protein band was excised after identification by Zn²⁺-based reverse staining of the gel (Hardy and Castellanos-Serra 2004), and the protein extracted by soaking crushed gel fragments in assay buffer (see "RNase P assay"). The concentration of *O. tauri* RPP was calculated using its extinction coefficient and absorbance at 280 nm.

Computational searches

The *O. tauri* chloroplast RPR gene was identified by BLAST (Altschul et al. 1990) in the chloroplast genome sequence (Palenik et al. 2007) querying with the sequence AGTCCG, which corresponds to the 5' half of the universally conserved helix P4. The list of hits was then examined manually for the presence of the 3' half (TCGGC) of the P4 helix within 200 nts downstream to AGTCCG, as has been observed with most bacterial RPRs.

The presence of homologs to experimentally validated eukaryal RPPs in the five sequenced prasinophyte genomes was analyzed at the DOE Joint Genome Institute server (<http://genome.jgi.doe.gov/>) using BLAST searches and the known sequences of yeast/human (eukaryotic nuclear) RPPs as queries.

Cloning of *O. tauri* organellar RPRs and pre-tRNA^{Leu}

To clone the RPRs, the encoding sequences were amplified by PCR from genomic DNA using gene-specific primers; a restriction site was engineered into each reverse primer (underlined sequence). Each PCR fragment was then digested with the corresponding restriction enzyme before being inserted downstream of the T7 promoter in pBT7 (Tsai et al. 2002) at the *Stu*I site and the restriction site introduced in the reverse primer. The primers for cloning *O. tauri* mitochondrial RPR were OtMtRPR-F (5'-GATAAGATGCCGCATGTCCG-3') and OtMtRPR-2R (5'-GACCCGGGATAAGATGCTTATAAGC-3'); and for *O. tauri* chloroplast RPR, OtCpRPR-F (5'-AAAAATAAAGAGTGGTTGCAG-3') and OtCpRPR-R (5'-GTGGATCCAAAAATAAAGTATAAGCCG-3'). Similarly, the gene encoding *O. tauri* mitochondrial pre-tRNA^{Leu} (GAG) was amplified with OtpLeu-F (5'-ATAATAATACATTACAGAATTTTTTTTGCAAAG-3') and OtpLeu-R (5'-ATGCATCCTGGTGCAAAGGTGGGACTTG-3'); the pre-tRNA^{Leu} PCR fragment was cloned into the *Stu*I site of pBT7. The sequence depicted in bold font in OtpLeu-F was added to create an extended 5'-leader sequence (from 21 to 29 nts) that would permit easy separation of RNase P-mediated cleavage products by gel electrophoresis. The cloned plasmids were confirmed by sequencing before further use.

In vitro transcription

Plasmids were first linearized by digesting pBT7-OtMtRPR with *Sma*I, pBT7-OtCpRPR with *Bam*HI, pBT7-OtMtpLeu and pUC19-NtCpGly (Raj and Gopalan, unpublished) with *Bst*NI, and pUC19-EcopTyr (Vioque et al. 1988) with *Fok*I. These plasmids then served as templates for T7 RNA polymerase-mediated run-off transcription using established protocols (Vioque et al. 1988). Templates encoding for the RPRs from *Synechocystis* PCC6803 and *Anabaena* PCC7120 were prepared as described previously (Pascual and Vioque 1999b). To obtain internally radiolabeled pre-tRNAs, these in vitro transcription reactions were supplemented with either ³²P-ATP or -GTP.

RNase P assays

The pre-tRNA processing assays were performed largely as described elsewhere (Vioque et al. 1988). *E. coli* RPR was pre-folded by first incubating in water at 50°C for 50 min, 37°C for 10 min, then in 1× assay buffer [10 mM HEPES (pH 7.5), 400 mM NH₄OAc, 10 mM Mg(OAc)₂, 5% (v/v) glycerol, 0.01% (v/v) IGEPAL] at 37°C for 30 min. The RNase P holoenzyme was reconstituted by incubating 1 nM folded RPR with either 10 nM *E. coli* RPP or *O. tauri* RPP in 1× assay buffer for 10 min at 37°C. Assays were then initiated by adding 100 nM pre-tRNA (a trace amount of which was internally labeled) and the incubation continued at 37°C. Reactions were quenched with an equal volume of 2× urea dye [7 M urea, 1 mM EDTA, 0.05% (w/v) xylene cyanol, 0.05% (w/v) bromophenol blue, and 10% (v/v) phenol] and separated on an 8% (w/v) polyacrylamide gel

containing 7 M urea. The extent of cleavage was visualized by phosphorimaging on the Typhoon (GE Healthcare) and quantitated by ImageQuant (GE Healthcare) to yield the initial velocity/turnover data.

For the single-turnover assays involving cyanobacterial RPRs and *O. tauri* RPP (Fig. 3), the RNase P holoenzyme was reconstituted with 50 nM RPR and 500 nM RPP and incubated with trace amounts of ³²P-labeled *E. coli* pre-tRNA^{Tyr} at 37°C for 45 min in 10 mM HEPES (pH 7.5), 50 mM MgCl₂. Reactions were quenched and analyzed as described above.

Single-turnover conditions were also employed for the assays with recombinant *O. tauri* PRORP (Fig. 7). Approximately 2 μM recombinant *O. tauri* PRORP was assayed for 30 min at 23°C as described elsewhere for *Arabidopsis* PRORP-1 (Gobert et al. 2010), with either internally labeled *O. tauri* pre-tRNA^{Leu} or tobacco chloroplast pre-tRNA^{Gly} (<1 nM) as the substrate. Reactions were quenched and analyzed as described above.

Genetic complementation

The function of *O. tauri* RPP was analyzed by genetic complementation of the *E. coli* thermosensitive *rnpA49* mutation (Schedl and Primakoff 1973). *E. coli* BL21(DE3)A49 (Guerrier-Takada et al. 1995) was transformed with pET-33b–OtrPP, which expresses His₆-tagged *O. tauri* RPP, and growth was assessed at 30°C (permissive) and 43°C (non-permissive). Cells transformed with pET-33b served as the negative control, while those transformed with pARE7, which express the wild type *E. coli* RPP (Vioque et al. 1988), were employed as the positive control. After obtaining the transformants at 30°C, individual colonies were grown overnight at 30°C in LB media with the following supplements: BL21(DE3)A49/pET-33b and BL21(DE3)A49/pET-33b–OtrPP, 4 μg/mL tetracycline + 50 μg/mL kanamycin; BL21(DE3)A49/pARE7, 4 μg/mL tetracycline + 100 μg/mL ampicillin. Tenfold serial dilutions of these cultures in fresh media were then plated on LB agar containing 4 μg/mL tetracycline, and growth was assessed after incubation at 30 or 43°C for 16 h.

Antibody generation and western blotting

Approximately 3 mg of affinity-purified *O. tauri* RPP was loaded on a preparative polyacrylamide gel, and the three *O. tauri* RPP fragments were excised and eluted together. To raise polyclonal antisera, 1 mg of such purified protein was injected subcutaneously into a white New Zealand rabbit. Antibodies were generated using standard procedures at the Animal Experimentation Facility, University of Sevilla. The antibodies were affinity purified (Harlow and Lane 1988) before use in western analysis. After separating 30 μg of total soluble *O. tauri* protein on a 12% (w/v) polyacrylamide-SDS gel, western blotting was performed using standard procedures. An anti-rabbit IgG-peroxidase (Sigma) was employed as the secondary antibody and the antigen detected using the SuperSignal West Femto Chemiluminescent substrate (Thermo Scientific).

RT-PCR

Total RNA was isolated from *O. tauri* RCC745 using either TriReagent LS (Sigma-Aldrich) or the RNeasy Plant Mini Kit (Qiagen) and treated with DNase I before proceeding. The *O. tauri* RPP transcript was reverse transcribed using OtC5rtPCR-R (5'-GTCACGGGTTGGTCGGC-3'; the sequence complementary to the termination codon is underlined) and BluePrint Reverse Transcriptase (Takara). Subsequently, with this RT reaction product as the template, PCR was performed with Ex Taq HS DNA Polymerase (Takara), and OtC5rtPCR-R + OtC5rtPCR-F (5'-GATGGCGTCGGGGGG-3'; the start codon is underlined). The *O. tauri*

was dialyzed into 30 mM HEPES (pH 7.6), 30 mM KCl, 6 mM MgCl₂. The concentration of *O. tauri* PRORP was calculated using its extinction coefficient and absorbance at 280 nm.

Results

Bacterial-like RPP in prasinophyte nuclear genomes

Since a bacterial RPP-like gene has not been reported in Eukarya, we inquired if it could be found in a basal, ancient clade such as Prasinophyceae. There is a hypothetical protein with significant homology to bacterial RPP in the nuclear genomes of *O. tauri* and four other prasinophyte algae: *Ostreococcus lucimarinus*, *Ostreococcus* sp. RCC809, *Micromonas pusilla*, and *Micromonas* sp. RCC299 (Fig. 1a). Although the bacterial RPP is typically <150 aa residues (e.g., *E. coli* RPP, 119 aa), the prasinophyte homologs are significantly larger (164–393 aa; Fig. 1a) and therefore have sequences/domains in addition to the bacterial RPP-like sequence. These additional domains show no significant similarity to other proteins or among themselves, except for a short homologous segment between *M. pusilla* and *M. sp. RCC299* (Fig. 1a). Further examination of these prasinophyte genes highlights other interesting features. First, in the two *Micromonas* RPPs, which are encoded by intron-containing sequences, the bacterial RPP-like domain is encoded by a single exon, suggesting exon shuffling as a possible mechanism by which this sequence was acquired. Second, there is an N-terminal appendage of 85 and 77 aa, respectively, to the bacterial RPP-like sequence in *O. tauri* and *Ostreococcus* sp. RCC809. Although some algorithms predict these N-terminal extensions to encode an organellar transit peptide, we are hesitant to draw a firm inference in this regard due to the low confidence in accurately predicting algal transit peptides.

Functional validation of the bacterial-like RPP in prasinophytes

The predicted coding sequence of *O. tauri* RPP was amplified by PCR and cloned into pET-33b, upstream of a sequence encoding the His6 tag to enable overexpression in *E. coli* BL21(DE3) cells and subsequent purification by immobilized metal (nickel)-affinity chromatography (IMAC). Three bands were detected in the affinity-purified fraction (Fig. 1b, left; ~34, 30 and 24 kDa). Although the expected size for *O. tauri* RPP-His6 is 27.6 kDa, we postulated that it corresponds to the 34 kDa band (the aberrant migration in SDS-PAGE likely due to its high isoelectric point of 10.5). The two smaller proteins might be N-terminally proteolyzed products of *O. tauri* RPP based on their co-purification with the putative full-length protein on IMAC and the fact that the His6 tag is at the C-terminus. Since the three proteins were readily detectable even in crude extracts of overexpressing cultures (before the start of purification), *O. tauri* RPP is probably subject to processing in *E. coli*. When the final purified preparation containing all three proteins was subjected to MALDI-TOF mass spectrometry, we determined that the longest protein (top band, T) has a size expected for the full-length protein lacking the first methionine, and the shorter proteins correspond to N-terminal deletions of 14 aa (middle band, M) or 66 aa (bottom band, B). Although the preparation with all three polypeptides could reconstitute with bacterial RPRs (see below), we gel-purified the three proteins individually (Fig. 1b, right) to ascertain if all three were functional.

We assayed each of these three recombinant *O. tauri* RPPs for their ability to reconstitute with the *E. coli* RPR to form a functional heterologous holoenzyme. Because *O. tauri* RPP appears to have an organellar transit peptide and might support a mitochondrial function, we cloned the mitochondrial pre-tRNA^{Leu}(GAG) from *O. tauri* and employed it as a substrate. Indeed, all three proteins could enhance cleavage of pre-tRNA^{Leu} by *E. coli* RPR; a tobacco chloroplast pre-tRNA^{Gly}(GCC) was also efficiently cleaved (data not shown). Under conditions optimal for *E. coli* RNase P activity, the holoenzyme assembled from *O.*

O. tauri RPP-T and *E. coli* RPR displayed a turnover number one-fifth of that observed with *E. coli* RNase P using pre-tRNA^{Leu} as the substrate (~12 vs. 56 min⁻¹; Fig. 2). Since the deletion derivatives of *O. tauri* RPP lacking the N-terminal 14 or 66 aa are ~70% as active as the full length (Fig. 2), the N-terminal 66 aa residues appear dispensable for pre-tRNA cleavage in vitro. These findings are not unexpected since a sequence alignment (Fig. 1a) reveals that the region of homology between *O. tauri* and *E. coli* RPP starts only at around residue 88 in the former.

We also performed reconstitution assays with the mixture of the three affinity-purified *O. tauri* RPPs and RPRs from two cyanobacteria, *Synechocystis* sp. PCC6803 and *Anabaena* sp. PCC7120 (Vioque 1992; Pascual and Vioque 1996). These heterologous reconstitutions are as active as the corresponding homologous reconstitutions under single-turnover conditions with an *E. coli* pre-tRNA^{Tyr} as substrate (Fig. 3).

Finally, we assessed if *O. tauri* RPP is functional in *E. coli*. For this purpose, we employed *E. coli* BL21(DE3)A49, a strain in which the R46H mutation in the RPP confers thermosensitivity (ts). Since growth is abolished at 43°C due to the inability of the RPP R46H mutant to support RNase P holoenzyme assembly, a plasmid that overexpresses a functional RPP can rescue this strain's ts phenotype (Jovanovic et al. 2002). We exploited this genetic complementation assay to investigate if the *O. tauri* RPP can functionally replace its *E. coli* counterpart. Indeed, this is the case as attested by the robust growth at 43°C of *E. coli* BL21(DE3)A49 transformed with pET-33b–OtRPP but not with the empty vector pET-33b (Fig. 4).

O. tauri RPP is expressed in vivo

While the heterologous reconstitution and in vivo complementation revealed the potential of *O. tauri* RPP to support bacterial RNase P catalysis, we sought to gain evidence for its expression in *O. tauri*. We isolated total RNA from *O. tauri* cells and amplified by RT-PCR the *O. tauri* RPP cDNA using gene-specific primers. An amplified product of the expected size was observed (Fig. 5a; 722 bp), whose identity was unambiguously established by DNA sequencing (data not shown). To confirm that *O. tauri* RPP is indeed synthesized in vivo, we first prepared antibodies against the purified *O. tauri* RPP (mixture of the three proteins; Fig. 1b) and performed western analysis of *O. tauri* cell extracts. We detected a protein of ~34 kDa (Fig. 5b), similar to the size expected for the full-length *O. tauri* RPP (compare Fig. 4b with 1b), an assertion further validated by western analysis of the recombinant versions (not shown).

Which RNA catalyst is aided by O. tauri bacterial-like RPP?

Our next objective was to investigate the possible RPRs whose function might either depend on or be aided by *O. tauri* RPP. Given the bacterial origins of organelles and the similarity of *O. tauri* RPP to bacterial RPP, we investigated whether *O. tauri* RPP could be a subunit of organellar RNase P.

While an RPR gene has already been annotated in the sequenced *O. tauri* mitochondrial genome (Fig. 6), we successfully identified a bacterial RPR-like gene in the chloroplast genome (Fig. 6; accession number CR954199, complementary strand positions 4,980 to 5,307). These findings are not surprising since genes coding for bacterial-like RPRs have been previously identified in the organellar genomes of *Nephroselmis olivacea*, another prasinophyte alga (Turmel et al. 1999a, b). Moreover, the organellar genomes in *O. tauri* and *N. olivacea* encode a near-complete suite of tRNAs, thus accounting for the need for RNase P activity in these organelles.

The *O. tauri* mitochondrial RPR is 269 nts long and has a 37% GC content, while the plastid RPR is 327 nts long with 29% GC content (Fig. 6a). The predicted secondary structures of these *O. tauri* RPRs (Fig. 6a)

agree with the bacterial consensus and contain all the universally conserved nucleotides (Chen and Pace 1997; Brown 1999; Marquez et al. 2005; Gopalan 2007). However, the secondary structure of these AU-rich *O. tauri* organellar RPRs, when compared to bacterial prototypes, show notable differences that likely account for the absence of crucial tertiary interactions essential for realizing pre-tRNA cleavage activity in vitro (McClain et al. 2010; Reiter et al. 2010).

The bacterial RPR, which is catalytically active in vitro, generates an active conformation through several GNRA tetraloop:helix docking tertiary contacts (L9:P1, L14:P8, L18:P8; L and P refer to loops and paired regions, respectively). In addition, pseudoknots (e.g., P6) aid the overall fold while other regions (e.g., L15) engage in sequence-specific substrate recognition (Reiter et al. 2010). Most of these attributes are absent in the *O. tauri* organellar RPRs. The *O. tauri* mitochondrial RPR is missing the P16/P17 helices and therefore also the pseudoknot generated by helix P6; it also lacks the P12 and P13 helices. Moreover, L9, typically a conserved, helix-stabilizing GNRA tetraloop in bacterial RPRs, is replaced with an AAGA tetraloop. The *O. tauri* plastid RPR lacks the P15 helix present in bacterial counterparts. In addition, L9 and L14 consist of UUG and ACCC in lieu of the typically conserved GNRA tetraloops at both locations that are involved in functionally important tertiary interactions (Pomeranz Krummel and Altman 1999). The GGU sequence in L15 of the bacterial RPR, which pairs with the 3'-terminal D73C74C75A76 motif in the pre-tRNA substrate (Kirsebom and Svard 1994; Kirsebom 2007), is also absent in both *O. tauri* organellar RPRs. Overall, while we appreciated that these structural differences in *O. tauri* organellar RPRs might undermine their ability to generate a native fold required for robust catalytic activity (in the absence of cognate protein factors), we did pursue pre-tRNA cleavage assays in vitro (see below).

Prior to embarking on biochemical studies, we inquired if the putative *O. tauri* organellar RPRs are expressed in vivo. RT-PCR using gene-specific oligonucleotides and total RNA as the template revealed that these two non-coding RNAs are indeed expressed in *O. tauri* (Fig. 6b). DNA sequencing of these RT-PCR products, which encompass most of the predicted RPR sequences, confirmed their identity. We assayed these *O. tauri* organellar RPRs under a number of different conditions (including different pre-tRNA substrates, temperature, pH, and range of divalent/monovalent ion concentrations) but have been unable to detect any pre-tRNA cleavage activity (data not shown). These RPRs also failed to form any functional holoenzymes with either *O. tauri* or bacterial RPPs (Fig. 3; data not shown). These findings parallel previous futile attempts with several organellar RPRs (Pascual and Vioque 1999a; De la Cruz and Vioque 2003)—the sole exception has been the weak enzymatic activity detected in the case of the plastid RPR from *Cyanophora paradoxa* (Li et al. 2007).

We can speculate on some reasons for the lack of activity of *O. tauri* organellar RPRs without and with *O. tauri* RPP. First, their low GC content (<37%) likely results in mis-folding/instability of in vitro transcribed preparations and therefore, a non-functional structure. Second, despite the resemblance of their secondary structure to other RPRs and the presence of universally conserved nucleotides believed to constitute the active site, these organellar RPRs lack some of the tertiary structure-stabilizing elements present in the bacterial RPR counterparts, which display robust activity. Third, the failure of *O. tauri* RPP to reconstitute with cognate organellar RPRs might reflect either the need for additional protein cofactors to realize RNase P activity or the possibility that it supports a function unrelated to tRNA maturation. These speculations can be verified only when the native organellar RNase P is purified and characterized.

PRORP in prasinophytes

O. tauri and other prasinophyte nuclear genomes encode a ~60 kDa protein homolog of *Arabidopsis* PRORP-1 (Gobert et al. 2010), which was shown to localize in mitochondria and chloroplasts and to catalyze pre-tRNA 5'-processing in vitro (under single-turnover conditions). The open reading frame (ORF) encoding *O. tauri* PRORP (~57 kDa) was amplified by PCR and cloned into pET-33b with a sequence encoding the His6 tag to enable overexpression in *E. coli* SHuffle T7 cells and purification to near homogeneity using tandem anion exchange chromatography and IMAC (Fig. 7a). This recombinant *O. tauri* PRORP promotes accurate in vitro 5'-processing of pre-tRNA^{Leu} (data not shown) and pre-tRNA^{Gly} (Fig. 7b). Although the activity we observed were in assays performed under single-turnover conditions (akin to the previous study on *Arabidopsis* PRORP-1), our result corroborates the finding of a single polypeptide capable of tRNA 5'-maturation and sets the stage for detailed kinetic studies to compare the catalytic efficiency of different PRORP- and RNP-based RNase P variants.

Discussion

An inventory of the RNase P variants in *O. tauri* (Fig. 8) highlights the striking diversity in this smallest free-living eukaryote. First, a bacterial RPP-like protein is encoded by the *O. tauri* nuclear genome. Although this protein is synthesized inside the cell and a recombinant version is functional with bacterial RPRs as assessed by in vitro reconstitution assays (Figs. 2, 3), further studies are needed to investigate if this protein in *O. tauri* is imported into mitochondria/plastid for subsequent assembly with the organellar RPRs (either alone or with additional protein cofactors). However, the size of the protein expressed in vivo is that expected for the full-length protein (Fig. 5), arguing against the presence of a cleavable transit peptide for import into organelles. It is not possible, however, to rule out aberrant migration arising from post-translational modifications of a shorter version. Second, the *O. tauri* mitochondrial and chloroplast genomes each encode a bacterial RPR-like RNA, which is expressed in vivo (Fig. 6). However, the exact functions of these RPRs remain to be determined. Third, we have now demonstrated that a recombinant version of *O. tauri* PRORP supports pre-tRNA 5'-processing in vitro (Fig. 7), although the subcellular locale of this protein in *O. tauri* has not been determined. Lastly, the full set of eukaryal (nuclear) RPPs was not found in any of the five prasinophyte algal genomes. Only POP5 and RPP30 could be identified with confidence in all five strains. In addition, there is no report of a putative *O. tauri* nuclear RPR (at least one resembling those found in eukaryotes ranging from yeast to human). In fact, the identity of plant/algal nuclear RPR remains unknown (Rosenblad et al. 2006).

Several interesting attributes and questions emerge from this growing appreciation of the variability of RNase P in *O. tauri* and, more broadly, in all domains of life. First, the discovery of a bacterial-like RPP in prasinophyte algae was unanticipated; a comprehensive phylogenetic study with additional taxa is required to investigate if this represents an instance of horizontal gene transfer from bacteria that co-habitat with *O. tauri*. In these prasinophyte proteins, the bacterial RPP-like domain is more similar to each other than to any bacterial RPP, suggesting that the acquisition of this domain happened once in a common prasinophyte ancestor. Our findings echo the recurring theme in recent studies on the surprising mixed ancestry of genes in some eukaryotes, especially phytoplankton. For instance, nitrogen transporters and related assimilation genes in *Micromonas* unexpectedly revealed that these genes share plant, algal or bacterial cousins (McDonald et al. 2010). It was postulated that the wide-ranging origins of certain genes might contribute to the fitness landscape of phytoplankton such as *Micromonas* and *Ostreococcus*, especially their ability to utilize energy sources that show seasonal variations in the oceanic habitat (McDonald et al. 2010). In a related vein, it is possible that the presence of multiple forms of RNase P (either RNP- or protein-based)

with distinct RNA processing traits affords a selective advantage in the specific ecological niche of these organisms.

Second, the diverse subunit make-up of RNase P in nature suggests multiple routes for generating its active site. What evolutionary forces and payoffs might underlie this pliability to employ a structurally diverse ensemble? For example, is it the high evolutionary rate of organelles, the plasticity of their genomes, and/or altered substrate specificity that has molded the extant organellar RNase P make-up? In this respect, it should be noted that several tRNA genes in *Ostreococcus* and *Micromonas* are permuted (Maruyama et al. 2010), and their processing by RNase P might require interactions of the enzyme with unusual structures in the permuted pre-tRNAs. The presence of diverse RNase P enzymes might afford a “divide and conquer” approach to deal with biogenesis/maturation of unusual pre-tRNA and non-coding RNA substrates.

Lastly, since the single plant organellar RNase P protein (PRORP, ~50 to 60 kDa) has homologs in evolutionarily distant eukaryotic lineages and can functionally substitute for the 135 kDa RNP complex in bacteria (Gobert et al. 2010), it is perplexing why cells, which now rely almost exclusively on proteins for catalytic and structural roles, have retained RNA-mediated catalysis at all in RNase P. Did retention of an RNA then confer advantages such as optimal recognition of RNA substrates through base pairing? Or is the RNP form better suited for enmeshing in a regulated network of macromolecular machineries mediating different aspects of gene expression (Reiner et al. 2006, 2008)? While the use of proteins and RNPs to perform the same biological task in different settings attests to nature’s versatility in designing catalysts, the daunting challenge is to understand the basis for these choices.

Acknowledgments

We thank Dr. M. L. S. Raj for cloning the tobacco chloroplast pre-tRNAGly(GCC) in pUC19. This work was supported by Ministerio de Ciencia e Innovación, Spain, and European Regional Fund [BFU2007-60651 (to A.V.)], Junta de Andalucía, Spain [P06-CVI-01692 (to A.V.)], European Union [ASSEMBLE grant agreement no. 227799 (to A.V.)], and the National Science Foundation [MCB-0238233 and MCB-0843543 (to V.G.)]. Pilar Bernal-Bayard was supported by a fellowship from Junta de Andalucía, Spain.

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
- Brown JW (1999) The ribonuclease P database. *Nucleic Acids Res* 27:314
- Chen JL, Pace NR (1997) Identification of the universally conserved core of ribonuclease P RNA. *RNA* 3:557–560
- Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res* 31:3497–3500
- De la Cruz J, Vioque A (2003) A structural and functional study of plastid RNAs homologous to catalytic bacterial RNase P RNA. *Gene* 321:47–56
- Derelle E, Ferraz C, Rombauts S, Rouze P, Worden AZ, Robbens S, Partensky F, Degroevé S, Echeynie S, Cooke R, Saeys Y, Wuyts J, Jabbari K, Bowler C, Panaud O, Piegue B, Ball SG, Ral JP, Bouget FY, Piganeau G, De Baets B, Picard A, Delseny M, Demaille J, Van de Peer Y, Moreau H (2006) Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc Natl Acad Sci USA* 103:11647–11652
- Esakova O, Krasilnikov AS (2010) Of proteins and RNA: the RNase P/MRP family. *RNA* 16:1725–1747
- Evans D, Marquez SM, Pace NR (2006) RNase P: interface of the RNA and protein worlds. *Trends Biochem Sci* 31:333–341
- Gobert A, Gutmann B, Taschner A, Gossringer M, Holzmann J, Hartmann RK, Rossmannith W, Giege P (2010) A single Arabidopsis organellar protein has RNase P activity. *Nat Struct Mol Biol* 17:740–744
- Gopalan V (2007) Uniformity amid diversity in RNase P. *Proc Natl Acad Sci USA* 104:2031–2032
- Guerrier-Takada C, Li Y, Altman S (1995) Artificial regulation of gene expression in *Escherichia coli* by RNase P. *Proc Nat Acad Sci USA* 92:11115–11119
- Hardy E, Castellanos-Serra LR (2004) “Reverse-staining” of biomolecules in electrophoresis gels: analytical and micropreparative applications. *Anal Biochem* 328:1–13
- Harlow E, Lane D (1988) *Antibodies: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- Holzmann J, Frank P, Löffler E, Bennett KL, Gerner C, Rossmannith W (2008) RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell* 135:462–474
- Jarrous N, Gopalan V (2010) Archaeal/eukaryal RNase P: subunits, functions and RNA diversification. *Nucleic Acids Res* 38:7885–7894
- Jovanovic M, Sanchez R, Altman S, Gopalan V (2002) Elucidation of structure-function relationships in the protein subunit of bacterial RNase P using a genetic complementation approach. *Nucleic Acids Res* 23:5065–5073

Keller MD, Selvin RC, Claus W, Guillard RRL (1987) Media for the culture of oceanic ultraphytoplankton. *J Phycol* 23:633–638

Kirsebom LA (2007) RNase P RNA mediated cleavage: substrate recognition and catalysis. *Biochimie* 89:1183–1194

Kirsebom LA, Svard SG (1994) Base pairing between *Escherichia coli* RNase P RNA and its substrate. *EMBO J* 13:4870–4876

Lai LB, Vioque A, Kirsebom LA, Gopalan V (2010) Unexpected diversity of RNase P, an ancient tRNA processing enzyme: challenges and prospects. *FEBS Lett* 584:287–296

Li D, Willkomm DK, Schön A, Hartmann RK (2007) RNase P of the *Cyanophora paradoxa* cyanelle: a plastid ribozyme. *Biochimie* 89:1528–1538

Liu F, Altman S (2010) Ribonuclease P. Springer-Verlag, New York

Marquez SM, Harris JK, Kelley ST, Brown JW, Dawson SC, Roberts EC, Pace NR (2005) Structural implications of novel diversity in eucaryal RNase P RNA. *RNA* 11:739–751

Maruyama S, Sugahara J, Kanai A, Nozaki H (2010) Permuted tRNA genes in the nuclear and nucleomorph genomes of photosynthetic eukaryotes. *Mol Biol Evol* 27:1070–1076

McClain WH, Lai LB, Gopalan V (2010) Trials, travails and triumphs: an account of RNA catalysis in RNase P. *J Mol Biol* 397:627–646

McDonald SM, Plant JN, Worden AZ (2010) The mixed lineage nature of nitrogen transport and assimilation in marine eukaryotic phytoplankton: a case study of *Micromonas*. *Mol Biol Evol* 27:2268–2283

Morales MJ, Dang YL, Lou YC, Sulo P, Martin NC (1992) A 105 kDa protein is required for yeast mitochondrial RNase P activity. *Proc Natl Acad Sci USA* 89:9875–9879

Palenik B, Grimwood J, Aerts A, Rouze P, Salamov A, Putnam N, Dupont C, Jorgensen R, Derelle E, Rombauts S, Zhou K, Otilar R, Merchant SS, Podell S, Gaasterland T, Napoli C, Gendler K, Manuell A, Tai V, Vallon O, Piganeau G, Jancek S, Heijde M, Jabbari K, Bowler C, Lohr M, Robbens S, Werner G, Dubchak I, Pazour GJ, Ren Q, Paulsen I, Delwiche C, Schmutz J, Rokhsar D, Van de Peer Y, Moreau H, Grigoriev IV (2007) The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proc Natl Acad Sci USA* 104:7705–7710

Pascual A, Vioque A (1996) Cloning, purification and characterization of the protein subunit of ribonuclease P from the cyanobacterium *Synechocystis* sp. PCC 6803. *Eur J Biochem* 241:17–24

Pascual A, Vioque A (1999a) Functional reconstitution of RNase P activity from a plastid RNA subunit and a cyanobacterial protein subunit. *FEBS Lett* 442:7–10

Pascual A, Vioque A (1999b) Substrate binding and catalysis by ribonuclease P from cyanobacteria and *Escherichia coli* are affected differently by the 3' terminal CCA in tRNA precursors. *Proc Natl Acad Sci USA* 96:6672–6677

Pomeranz Krummel DA, Altman S (1999) Verification of phylogenetic predictions in vivo and the importance of the tetraloop motif in a catalytic RNA. *Proc Natl Acad Sci USA* 96:11200–11205

Reiner R, Ben-Asouli Y, Krilovetzky I, Jarrous N (2006) A role for the catalytic ribonucleoprotein RNase P in RNA polymerase III transcription. *Genes Dev* 20:1621–1635

Reiner R, Krasnov-Yoeli N, Dehtiar Y, Jarrous N (2008) Function and assembly of a chromatin-associated RNase P that is required for efficient transcription by RNA polymerase I. *PLoS One* 3:e4072

Reiter NJ, Osterman A, Torres-Larios A, Swinger KK, Pan T, Mondragon A (2010) Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature* 468:784–789

Rosenblad MA, Lopez MD, Piccinelli P, Samuelsson T (2006) Inventory and analysis of the protein subunits of the ribonucleases P and MRP provides further evidence of homology between the yeast and human enzymes. *Nucleic Acids Res* 34:5145–5156

Schedl P, Primakoff P (1973) Mutants of *Escherichia coli* thermosensitive for the synthesis of transfer RNA. *Proc Natl Acad Sci USA* 70:2091–2095

Tsai HY, Lai LB, Gopalan V (2002) A modified pBluescript-based vector for facile cloning and transcription of RNAs. *Anal Biochem* 303:214–217

Turmel M, Lemieux C, Burger G, Lang BF, Otis C, Plante I, Gray MW (1999a) The complete mitochondrial DNA sequences of *Nephroselmis olivacea* and *Pedinomonas minor*. Two radically different evolutionary patterns within green algae. *Plant Cell* 11:1717–1730

Turmel M, Otis C, Lemieux C (1999b) The complete chloroplast DNA sequence of the green alga *Nephroselmis olivacea*: insights into the architecture of ancestral chloroplast genomes. *Proc Natl Acad Sci USA* 96:10248–10253

Vioque A (1992) Analysis of the gene encoding the RNA subunit of ribonuclease P from cyanobacteria. *Nucleic Acids Res* 20:6331–6337

Vioque A, Arnez J, Altman S (1988) Protein-RNA interactions in the RNase P holoenzyme from *Escherichia coli*. *J Mol Biol* 202:835–848

Walker SC, Engelke DR (2006) Ribonuclease P: the evolution of an ancient RNA enzyme. *Crit Rev Biochem Mol Biol* 41:77–102

Wang G, Chen HW, Oktay Y, Zhang J, Allen EL, Smith GM, Fan KC, Hong JS, French SW, McCaffery JM, Lightowlers RN, Morse HC 3rd, Koehler CM, Teitell MA (2010) PNPASE regulates RNA import into mitochondria. *Cell* 142:456–467

Worden AZ, Lee JH, Mock T, Rouze P, Simmons MP, Aerts AL, Allen AE, Cuvelier ML, Derelle E, Everett MV, Foulon E, Grimwood J, Gundlach H, Henrissat B, Napoli C, McDonald SM, Parker MS, Rombauts S, Salamov A, Von Dassow P, Badger JH, Coutinho PM, Demir E, Dubchak I, Gentemann C, Eikrem W, Gready JE, John U, Lanier W, Lindquist EA, Lucas S, Mayer KF, Moreau H, Not F, Otilar R, Panaud O, Pangilinan J, Paulsen I, Piegu B, Poliakov A, Robbens S, Schmutz J, Toulza E, Wyss T, Zelensky A, Zhou K, Armbrust EV, Bhattacharya D, Goodenough UW, Van de Peer Y, Grigoriev IV (2009) Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science* 324:268–272

Figure captions

Figure 1. Bacterial-like RPP in Prasinophyceae. a Alignment of the *E. coli* RPP sequence with homologs from several prasinophyte genomes. Homologous proteins were identified in the nuclear genome sequence of *Micromonas pusilla* CCMP1545 (*M. pusilla*, XM_003057480), *Micromonas* sp. RCC299 (*M. RCC299*, XM_002501961), *Ostreococcus lucimarinus* (*O. lucimarinus*, XM_001421023), *Ostreococcus tauri* (*O. tauri*, XM_003082501), and *Ostreococcus* sp. RCC809 (*O. RCC809*, the Joint Genome Institute Database, <http://genome.jgi-psf.org/>). The alignment was generated with Clustal W (Chenna et al. 2003) and manually refined. The regions that share high similarity with *E. coli* RPP are boxed. While positions identical in all five prasinophyte proteins are indicated with asterisks above the alignment, those shared with *E. coli* RPP are indicated with asterisks below the alignment. The black triangles show the position of predicted introns (according to the JGI annotation) in the two *Micromonas* strains. Underlined is a CCHC zinc-finger motif in *M. pusilla* RPP; the dotted-line box highlights a segment of similarity between *M. pusilla* and *M. RCC299* RPPs outside the *E. coli* RPP-like region. b Purification of recombinant *O. tauri* RPP (OtrPP). Coomassie blue-stained polyacrylamide gels illustrate purity of the OtrPP-containing peak fraction from nickel affinity chromatography (left) and of the individual isolated polypeptides (right). The predicted N-terminus of each polypeptide (T top, M middle, B bottom), as determined by MALDI-TOF mass spectrometry, is indicated in (a)

Figure 2. Heterologous reconstitution of RNase P activity with *O. tauri* RPP and *E. coli* RPR. RNase P holoenzymes were assembled using *E. coli* RPR (EcRPR) and either purified *E. coli* RPP (EcRPP) or each gel-purified polypeptide of *O. tauri* RPP (OtrPP-T, -M, or -B). The first lane is a negative control with EcRPR in the absence of any protein cofactor, and the “– EcRPR” lanes are controls performed with only the indicated proteins but no RPR. The multiple-turnover reaction rate for each of the four reconstituted holoenzymes is indicated at the bottom of the respective time courses

Figure 3. Heterologous reconstitution of RNase P activity with *O. tauri* RPP and cyanobacterial RPR. Reconstitution of RNase P activity with purified recombinant RPP and in vitro transcribed RPR from *Synechocystis* PCC6803 (Sy), *Anabaena* PCC7120 (An), and *O. tauri* mitochondria (Ot). Assays were performed under single-turnover conditions (see “Materials and methods” for details). The first lane is a negative control without RPR or RPP, and the “– RPR” lanes are controls performed with only the indicated RPP but no RPR

Figure 4. *O. tauri* RPP functionally substitutes for a bacterial RPP in vivo. Tenfold serial dilutions of the temperature-sensitive *E. coli* BL21(DE3)A49 strain harboring pARE7 (expressing *E. coli* RPP), the empty vector pET-33b, or pET-33b–OtrPP (expressing *O. tauri* RPP) were assessed for growth on LB agar plates at 30°C (permissive) or 43°C (non-permissive) for 16 h

Figure 5. Expression of the bacterial-like RPP in *O. tauri*. a Detection of the 722 bp cDNA encoding the *O. tauri* RPP by RT-PCR (+), but not by PCR alone (–), in total RNA isolated from *O. tauri* cells. b Western analysis of *O. tauri* (Ot) total protein extract with polyclonal antibodies raised against *O. tauri* RPP

Figure 6. RPRs from *O. tauri* mitochondria and chloroplasts. a Secondary structure models of RPRs from *O. tauri* mitochondria (mt; left) and chloroplast (cp; right). Paired regions (helices) are labeled as P1, P2, etc., consecutively from 5' to 3' following the nomenclature used for bacterial RPRs. The prefix L refers to loops

capping paired regions. Universally conserved nucleotides (Chen and Pace 1997; Brown 1999; Marquez et al. 2005; Gopalan 2007) are highlighted with a black circle. Solid lines show long-distance interactions that generate pseudoknots P4 and P6, as well as other tertiary interactions. b Expression of organellar RPRs in *O. tauri*. Detection of the cDNAs corresponding to the RPR from *O. tauri* mitochondria (left; 255 bp) and chloroplast (right; 304 bp) by RT-PCR (+) but not by PCR alone (-) in total RNA isolated from *O. tauri* cells

Figure 7. Purification and activity of recombinant *O. tauri* PRORP. a Coomassie blue-stained polyacrylamide gel illustrates purity of the PRORP-containing peak fraction from nickel-affinity chromatography. b Pre-tRNA 5'-processing activity of recombinant *O. tauri* PRORP. The first lane (-) is a negative control without any enzyme, the second lane (+) is a positive control with *E. coli* RNase P, and the last lane (PRORP) is a reaction with *O. tauri* PRORP

Figure 8. The potential complexity and diversity of RNase P in *O. tauri*. Schematic summarizing the RNase P components encoded by each of the three *O. tauri* genomes. They were identified based on sequence homology

Figure 2

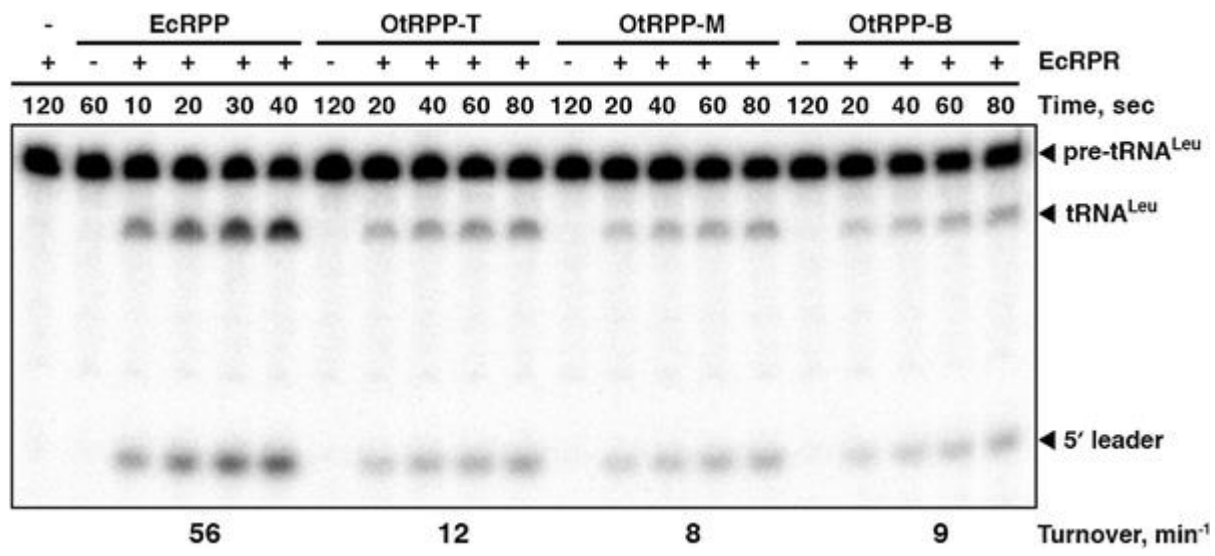


Figure 3

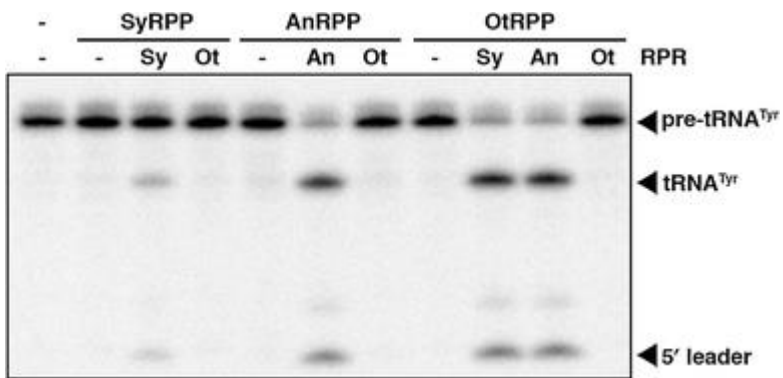


Figure 4

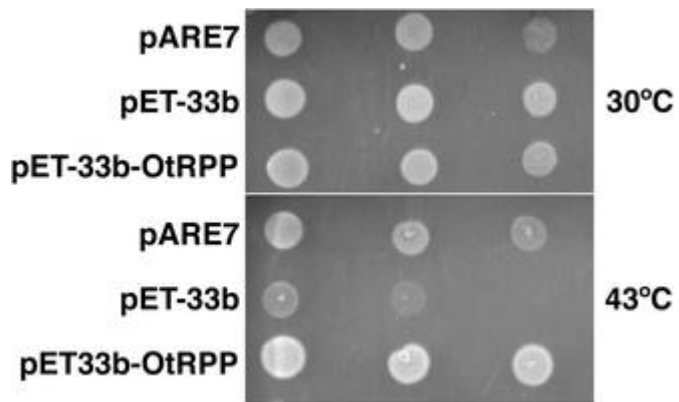


Figure 5

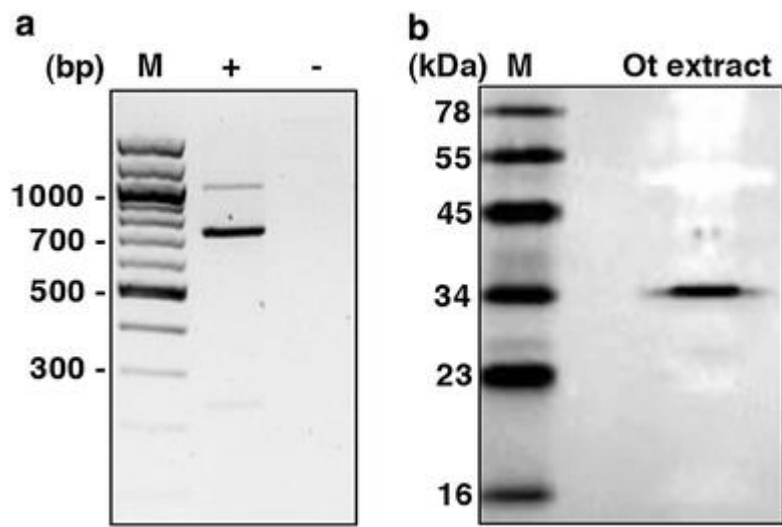
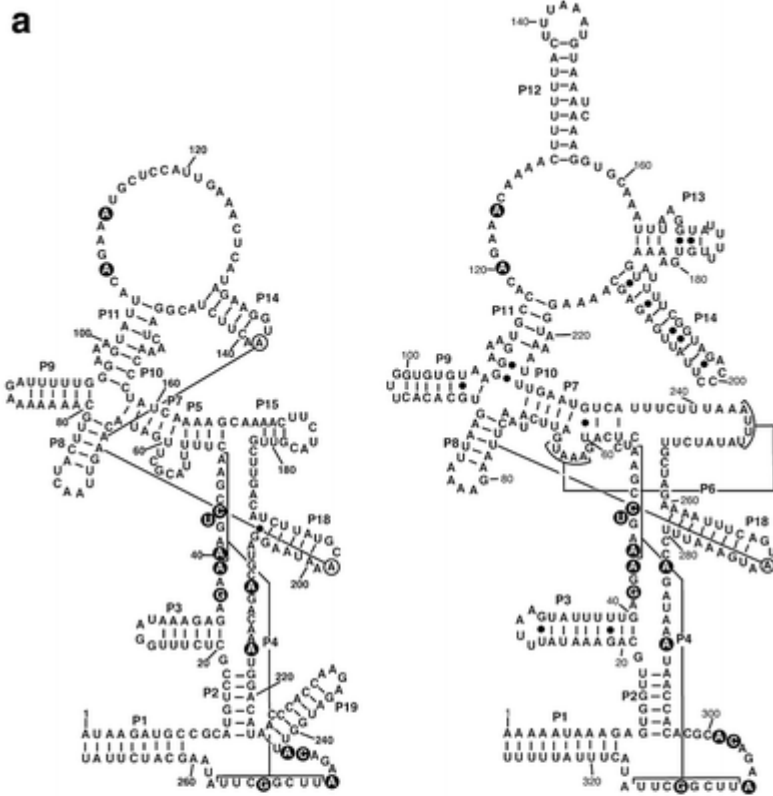
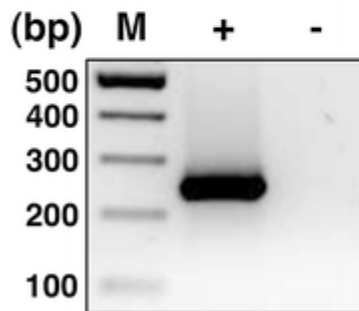


Figure 6



b *O. tauri* mt RPR



O. tauri cp RPR

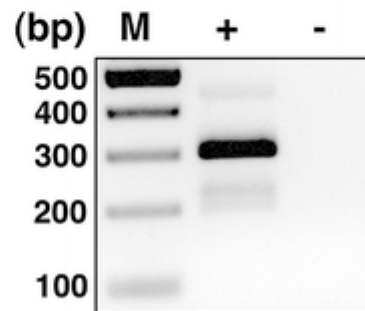


Figure 7

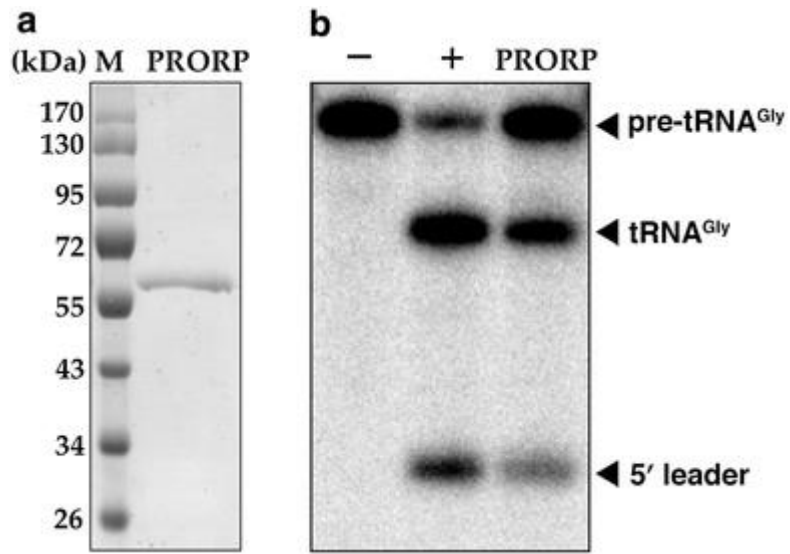


Figure 8

