

ARTICLE

Received 6 Jul 2012 | Accepted 5 Dec 2012 | Published 15 Jan 2013

DOI: 10.1038/ncomms2358

OPEN

Structural insights into protein-only RNase P complexed with tRNA

Anthony Gobert^{1,*}, Franziska Pinker^{1,2,*}, Olivier Fuchsbauer², Bernard Gutmann¹, René Boutin³, Pierre Roblin^{4,5}, Claude Sauter² & Philippe Giegé¹

RNase P is the essential activity removing 5'-leader sequences from transfer RNA precursors. RNase P was always associated with ribonucleoprotein complexes before the discovery of protein-only RNase P enzymes called PRORPs (PROteinaceous RNase P) in eukaryotes. Here we provide biophysical and functional data to understand the mode of action of PRORP enzymes. Activity assays and footprinting experiments show that the anticodon domain of transfer RNA is dispensable, whereas individual residues in D and T ψ C loops are essential for PRORP function. PRORP proteins are characterized in solution and a molecular envelope is derived from small-angle X-ray scattering. Conserved residues are shown to be involved in the binding of one zinc atom to PRORP. These results facilitate the elaboration of a model of the PRORP/transfer RNA interaction. The comparison with the ribonucleoprotein RNase P/transfer RNA complex suggests that transfer RNA recognition by PRORP proteins is similar to that by ribonucleoprotein RNase P.

¹Institut de Biologie Moléculaire des Plantes du CNRS, Université de Strasbourg, 12 rue du Général Zimmer, 67084 Strasbourg, France. ²Institut de Biologie Moléculaire et Cellulaire du CNRS, Architecture et Réactivité de l'ARN, Université de Strasbourg, 15 rue René Descartes, 67084 Strasbourg, France.

³Laboratoire d'Hydrologie et de Géochimie du CNRS, 1, rue Blessig, 67084 Strasbourg, France. ⁴Synchrotron SOLEIL, l'Orme des Merisiers Saint-Aubin, 91410 Gif-sur-Yvette, France. ⁵URBIA-Nantes, INRA Centre de Nantes, 60 rue de la Géraudière, 44316 Nantes, France. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to P.G. (email: philippe.giege@ibmp-cnrs.unistra.fr) or to C.S. (email: c.sauter@ibmc-cnrs.unistra.fr).

RNase P is the ubiquitous activity that catalyses the 5'-maturation of transfer RNAs (tRNAs), as well as of a number of other substrates such as ribosomal RNA, messenger RNA, transfer-messenger RNA or riboswitches¹⁻³. RNase P was first described in bacteria where it is composed by a ribonucleoprotein (RNP) complex whose RNA component (P RNA) holds the catalytic activity⁴. RNP RNase P was later found in all three main branches of life, that is, Bacteria, Archaea and Eukarya, and was thus believed to occur universally as a RNP complex⁵. This concept was challenged by early experiments in human mitochondria and spinach chloroplasts that suggested that another type of RNase P devoid of RNA component existed in these organelles^{6,7}. Still the dogma of the universality of RNP RNase P remained until the recent characterization of a novel type of RNase P in human mitochondria and plant organelles^{8,9}. This novel variant is composed of a single protein that we called PRORP (for PROteinaceous RNase P) and occurs in nearly all major phyla of eukaryotes⁹. Furthermore, RNP RNase P has not been retained in all organisms because, in both *Arabidopsis* and *Trypanosoma*, PRORP enzymes were found to support RNase P activity in both organelles and the nucleus^{10,11}.

The discovery of PRORP enzymes leads to the question of the respective mode of action of RNP and protein enzymes catalysing the same reaction. RNP RNase P activity is well characterised^{2,12}, in particular, recent advances such as the determination of the three-dimensional structure of a bacterial RNase P in complex with tRNA have been very important developments¹³. Substrate recognition by RNP RNase P involves the binding to regions distant from the actual cleavage site. It includes stacking interactions between bases in the D and T Ψ C loops of tRNAs and the P RNA specificity (S) domain, an A-minor interaction at the acceptor stem and the formation of canonical base pairs at the 3'-end of tRNA. In particular, key interactions take place between the unstacked bases G19 and C56 of tRNA and the S domain of RNase P. It is also notable that no interaction takes place with the anticodon arm of tRNA. The catalytic active site of RNP RNase P is composed of phosphate backbone moieties, a conserved uridine and at least two catalytically important metal ions¹³.

In contrast, the mode of action of PRORP proteins is unknown. These protein-only RNase P enzymes are characterized by the presence of pentatricopeptide (PPR) repeats¹⁴ in their N-terminal part that are believed to be involved in RNA binding and possess an upstream zinc-finger-like motif. The most conserved part of PRORP enzymes lies in their C-terminal part. This region was predicted to be a metallo-nuclease domain⁸ and consigns PRORP to the large family of PIN-like/NYN (N4BP1, YacP-like Nuclease) domain putative ribonucleases¹⁵.

Initial comparison of PRORP and RNP RNase P has suggested that the two classes of enzymes share common features. They both appear to require Mg²⁺ for phosphodiester hydrolysis and both generate 5'-phosphate and 3'-hydroxyl products⁸. However, studies using spinach chloroplast extracts and recombinant *Arabidopsis* PRORP have shown that PRORP is a fundamentally different catalyst than RNP RNase P. The replacement of the phosphodiester backbone of a precursor tRNA by a phosphorothioate moiety at the level of the 5'-maturation site resulted in a strong inhibition of bacterial RNase P activity, while PRORP activity was unaltered^{16,17}.

To gain functional insight into this novel type of RNase P activity, we investigated how *Arabidopsis* PRORP1 binds tRNA substrates and we performed a biophysical characterization of PRORP1 and 2. This enabled us to define initial mechanistic data on PRORP mode of action. The proposed mode of RNA recognition by PRORP shows striking similarity with that of RNP RNase P, which suggests that protein-only RNase P might

have converged to the same tRNA-binding strategy as RNP RNase P.

Results

tRNA cis elements required for PRORP activity. To get mechanistic insights into the mode of action, in particular of RNA recognition of PRORP enzymes, we performed RNase P cleavage assays with recombinant PRORP1 and different mutants of mitochondrial tRNA^{Cys} precursor, a known substrate of PRORP1 *in vivo* (Fig. 1). We first removed the anticodon domain from the tRNA precursor, which did not result in significant decrease of cleavage by PRORP. Then, the removal of both the D and anticodon domains was tested. The resulting mini helix was not cleavable by PRORP. As PRORP enzymes are able to cleave the 5'-leader sequence of any tRNA of canonical structure *in vitro*⁸⁻¹¹, we postulated that the determinants for tRNA recognition by PRORP must reside among positions universally conserved in tRNAs¹⁸. We thus applied point mutations to such positions in tRNAs to investigate their effect on PRORP activity. The mutation of G18 in the D-loop to A and C, respectively, resulted in severe impairment and total loss of RNase P activity, whereas the mutation of G19 to A or C did not affect RNase P activity. However, mutations of C56 in the T Ψ C loop to A or G resulted in total loss of RNase P cleavage. In the same loop, mutations of G57 to A and C resulted in unaffected and total loss of PRORP cleavage, respectively, consistent with the conservation of a purine at position 57 in tRNA¹⁸. Next, the exchange of G-C, the first base pair of the acceptor stem by C-G did not result in decreased cleavage efficiency, although we cannot exclude that mis-cleavage did not occur. Finally, we investigated the nature of the 3'-end of tRNA precursors. The absence of a 3'-trailer sequence did not affect 5'-cleavage, whereas the occurrence of a 3'-CCA group strongly reduced RNase P activity (Fig. 1). Further analyses will be necessary to determine if 3'-CCA groups act as PRORP binding antideterminants for all tRNAs, and to uncover the precise involvement of the length and the nature of residues in 3'-trailer sequences for PRORP activity.

Taken together, our results show that the anticodon domain is not involved in RNA recognition by PRORP. The nature of residues at positions 1 and 72 is not discriminant for the activity, while the 3'-CCA seems to act as an antideterminant for PRORP binding. As residues at positions 18 and 57 are involved in interactions between loops D and T Ψ C of tRNA¹⁸, an interaction between loops D and T Ψ C appears to be strictly required for PRORP function. However, the conservation of the G19:C56 interaction does not appear to be critical, although the presence of a cytidine at position 56 seems to be indispensable. Thus, precise residues in loops D and T Ψ C seem to be essential for substrate recognition by PRORP.

tRNA residues in interaction with PRORP. We performed a footprinting analysis in order to map precisely contact points between a PRORP protein and its tRNA precursor substrate (Fig. 2). To determine the tRNA regions that are in interaction with PRORP1, a mitochondrial tRNA^{Cys} precursor containing a leader sequence of five nucleotides was incubated either alone or in complex with PRORP1 and subjected to digestion by nucleases. As the PRORP/tRNA complex used here was UV-crosslinked, we verified with an *in vitro* activity assay that substrate binding had resulted in a catalytically active complex (Supplementary Fig. S1). Three different nucleases were used for this analysis, RNase V1 that only cleaves base-paired RNA regions, RNase T1 that cuts single-stranded RNA only after guanines and RNase A that cleaves single-stranded RNA after cytidines and uridines. Positions of residues protected from RNase digestion by PRORP

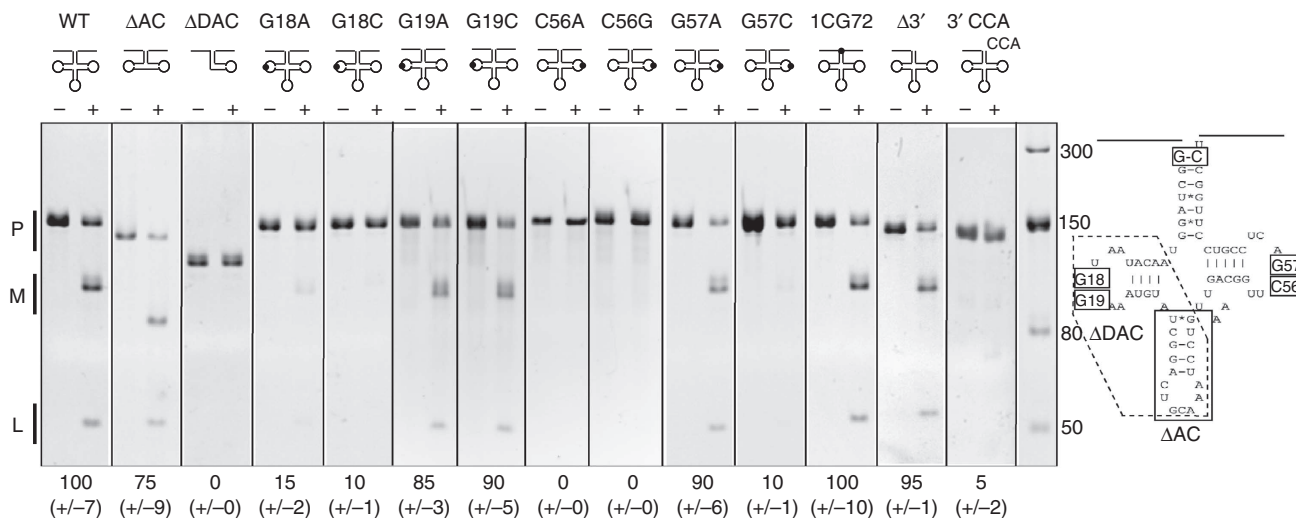


Figure 1 | RNase P *in vitro* cleavage assays performed with *Arabidopsis* recombinant PRORP1 and variants of mitochondrial tRNA^{Cys} precursors.

+ and - indicate the absence and presence of PRORP proteins in the reactions. WT is the wild-type tRNA^{Cys}, ΔAC the tRNA without the anticodon domain, and ΔDAC without both anticodon and D domains. G18A, for example, shows a tRNA where the guanosine at position 18 was mutated into an adenine. 1CG72 is a tRNA where the G-C base pair at positions 1 and 72 was swapped to a C-G. Δ3'-shows a tRNA precursor without 3'-trailer sequence and 3'-CCA, the precursor with a mature 3'-end containing a CCA. P stands for tRNA precursors, M the 5'-mature products and L the cleaved 5'-leader fragments. The molecular weights of markers are given in nucleotides. PRORP cleavage products were quantified with ImageGauge (Fujifilm). Values were normalized so that 100 corresponds to the cleavage efficiency observed for wild-type tRNA^{Cys} precursor. Cleavage efficiencies are given below the respective panels together with s.d.'s for three representative experiments.

could be mapped down to individual nucleotides through the comparison of RNase digestion profiles with an RNase T1 ladder and the alkaline hydrolysis profile of the mitochondrial tRNA^{Cys} precursor. We observed that discrete positions in the tRNA D-loop, namely U16, G18 and G19 were protected from nuclease digestion by PRORP. Similarly, C56 in the TψC loop was protected from nuclease digestion by PRORP. No other position, in particular close to the actual cleavage site of the tRNA precursor could be reproducibly identified as a site of nuclease protection by PRORP. Altogether, this indicates that individual residues that are in close spatial vicinity in loops D and TψC of tRNA are binding sites for protein-only RNase P enzymes (Fig. 2).

Structural properties of PRORP 1 and 2 in solution. We characterized in parallel both organellar and nuclear enzymes and focussed our study on PRORP1 and PRORP2 (which displays 80% sequence identity with PRORP3). *Arabidopsis* PRORP proteins are active as single-protein enzymes⁹. Their hydrodynamic properties in size exclusion chromatography and in dynamic light scattering confirmed that they are monomers in solution (Fig. 3). The molecular mass determined for PRORP2 in multi-angle light scattering is 62 kDa in good agreement with that calculated from the sequence (60 kDa).

These monodisperse PRORP samples led to sharp synchrotron radiation circular dichroism (SRCD) spectra (Fig. 3d) indicating that both PRORP1 and PRORP2 have a high content in α-helices. The evaluation of PRORP secondary structure content indicates 36/39% of α-helices, 15/16% of β-strands in PRORP1 / PRORP2, respectively. This observation is consistent with structure predictions based on sequence analysis⁹(Supplementary Fig. S2).

PRORP samples were further studied by small-angle X-ray scattering (SAXS), a method of structural characterization providing information on the size and shape of biological macromolecules in solution^{19–21}. In these experiments, the two enzymes produced very similar scattering curves at small angles, and their estimated gyration radius is $R_g = 33 \text{ \AA}$ corresponding

to a monomer in solution (Fig. 4). An experimental setup that allows the SAXS analysis downstream of a gel-filtration separation enabled the acquisition of scattering data for PRORP2 with lower noise at higher angles (Fig. 4, blue plot). The derived P(R) function that evaluates the distribution of distances inside the molecular object, confirms the value of R_g and is compatible with an object made of two structured domains. The tail of the distribution ($80 < r < 110 \text{ \AA}$) suggests the presence of extension(s), which may correspond to either N-terminal or C-terminal regions. PRORP proteins are slightly more compact in solution than archaeal and bacterial RNase P RNAs, which display R_g and d_{max} of 38–48 Å and 120–190 Å in SAXS, respectively^{22,23}. Overall, they appear as monomeric enzymes with two-domains essentially made of α-helices.

PRORP proteins are zinc-binding enzymes. The analysis of PRORP sequence conservation across eukaryotes revealed that a certain number of residues are highly conserved throughout evolution and might thus be of functional importance⁹. Among them, a putative zinc-finger-like structure is split in two separate motifs. The first motif (CxxC) contains two conserved cysteines upstream of the NYN domain at positions 344 and 347 for PRORP1 (281 and 284 for PRORP2), whereas the second motif involves a conserved histidine and a cysteine, downstream of the NYN domain, at positions 548 and 565, respectively, (Fig. 3a) (494 and 511 for PRORP2). These four particular residues were chosen as best candidates to form a zinc-binding pocket as no other cysteine or histidine outside the catalytic NYN domain is highly conserved among PRORP sequences. We used inductively coupled plasma mass spectrometry to investigate the association of metal cofactors to PRORP proteins. Zinc (⁶⁶Zn) was present at $29.49 \pm 1.53 \text{ p.p.b.}$ in a $30\text{-}\mu\text{g ml}^{-1}$ PRORP1 solution. This corresponds to the occurrence of one zinc atom per PRORP molecule. Other metals were only found as traces. To investigate the importance of the conserved residues for zinc binding, we mutated the four residues to alanines, expressed and purified to

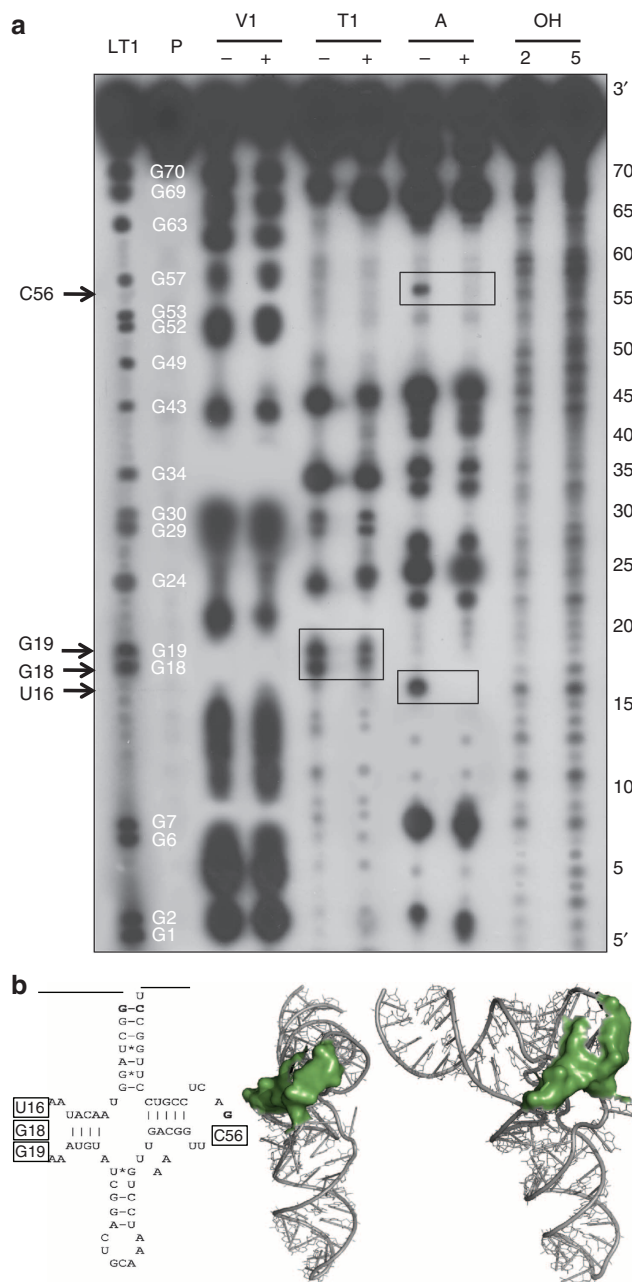


Figure 2 | Footprinting analysis of mitochondrial tRNA^{Cys} precursor in complex with PRORP1. (a) Samples were subjected to partial RNase V1, RNase T1 and RNase A digestions. + and – mean that PRORP proteins were present or absent in the reactions. P represents the tRNA precursor probe. LT1 shows an RNase T1 ladder with the corresponding positions of Gs in the tRNA sequence indicated in white. OH show alkaline hydrolysis of the tRNA probe performed for 2 and 5 min to generate an RNA ladder with single-nucleotide increments. RNA samples were separated by high resolution denaturing PAGE. tRNA positions were precisely mapped with the T1 and alkaline ladders. Boxed positions, also indicated on the left by arrows, correspond to tRNA positions reproducibly found protected from nuclease treatment by PRORP interaction in three replicate experiments. (b) Secondary and tertiary structural model of mitochondrial tRNA^{Cys} with boxes and green surfaces indicating residues protected by PRORP in footprinting experiments.

homogeneity the respective PRORP1 mutants. The analysis of zinc content in the mutants revealed that in the C344 and C347 mutants zinc levels were reduced by 19% and 29%, respectively,

whereas the H548 and the C565 mutants zinc levels decreased by 60% and 75%, respectively (Table 1). As a control we analysed zinc content in the DD474–475 catalytic mutant⁹ and found that it was similar to that of wild-type PRORP. The increased lability of zinc in the cysteine and histidine PRORP mutants suggests that the four residues are involved in the stable binding of zinc and that the downstream conserved motif has a stronger affinity for the metal than the upstream CxxC coordination element. We also analysed the capacity of the mutant proteins to perform RNase P activity. The single C565 mutant protein had impaired RNase P activity (Supplementary Fig. S3). The highest lability of zinc in this mutant might have resulted in an unstable protein fold, thus affecting its activity.

Model of the PRORP/tRNA complex. Structural models of PRORP were generated by homology modelling as implemented on the Phyre server²⁴. The structure prediction was limited to the two main domains (PPR and NYN) for which templates were identified with Phyre2. Best hits for the three PRORP sequences were a TPR domain (PDBid: 2o0e) and the catalytic domain of an RNase (PDB id: 3v32), respectively (see Supplementary Fig. S2). We also established the structure model of the *Arabidopsis* tRNA^{Cys} used to illustrate footprinting experiments and activity assays. The latter data were combined with the SAXS envelope in order to position domains of PRORP2 (for which we had best SAXS data) with respect to the tRNA substrate. In the docking process, the N-terminal RNA recognition module of PRORP2 containing PPR repeats⁹ was placed next to loops D and TψC of the folded tRNA, in particular in contact with positions U16, G18, G19 and C56. For the C-terminal part of the protein, the two conserved aspartates at positions 474 and 475 (Fig. 3a) that were shown to be part of the catalytic active site of PRORP⁹ were placed in close vicinity of the tRNA +1 position, where RNase P cleavage takes place (Fig. 5). Our model highlights notable similarities in tRNA-binding mode with the complex of bacterial RNP RNase P where the specificity domain of RNase P RNA interacts with the residues G19 and C56 of the tRNA (Fig. 5).

Discussion

PRORP enzymes were identified as members of the PPR family, a huge class of RNA-binding proteins ubiquitous in eukaryotes²⁵. These proteins can be divided into two main super-groups (P and PLS) according to the occurrence of specific classes of PPR domains and of additional C-terminal domains. Several lines of evidence suggest that both types of P and PLS proteins recognize primary sequences of RNA¹⁴. Interestingly PRORP enzymes do not belong to the two established super-groups. With only very few canonical PPR domains, and the presence of non-canonical putative PPR repeats, they rather define a new subfamily of PPR proteins. As PRORP enzymes bind any tRNA of canonical structure, it is possible that PRORP proteins recognize structured elements of RNA and thus have a mode of RNA recognition distinct from other PPR proteins. Alternatively, our favoured hypothesis is that PPR repeats in PRORP might specifically recognize individual nucleotides in tRNA loops, in particular unstacked bases or residues not involved in Watson–Crick interactions, which are highly conserved among tRNA sequences.

The biophysical characterization of PRORP enzymes has validated bioinformatic predictions and enabled to build a model of the active enzymatic complex. The predominance of SRCD signal for α -helices (Fig. 3d) is in agreement with fold recognition predictions, PPR repeats and NYN domains being mostly composed of α -helices^{15,26}. Although the N-terminal region of PRORP does not contain >2–3 canonical PPR motifs, the presence of non-canonical putative PPR domains suggests

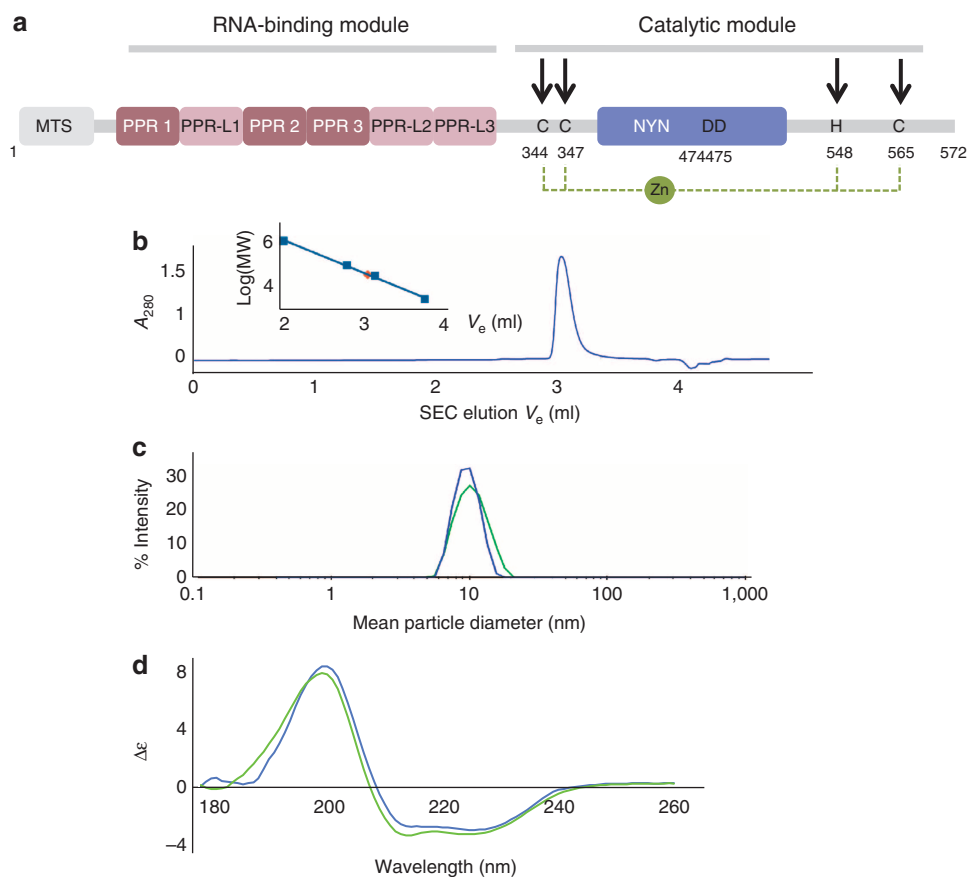


Figure 3 | Characterization of PRORP proteins in solution (a) Organization along the sequence of PRORP proteins. In the RNA-binding domain PPR and PPR-L show canonical PPR repeats and putative PPR-like motifs, respectively. For PRORP1, aspartates at positions 474 and 475 are in the catalytic pocket of the enzyme⁹, whereas cysteines and a histidine at positions 344, 347, 548 and 565 (indicated by black arrows) are proposed to form a zinc-binding pocket (dashed line) that could stabilize the catalytic domain of PRORP. (b) Analysis of PRORP2 oligomeric state in solution by analytical gel-filtration (BioSEC3 column) leading to a MW estimation of 72 kDa (red diamond: $\text{Log}(\text{MW}) = 4.85$) by comparison with the elution of model proteins (thyroglobulin: 660 kDa; BSA monomer and dimer: 66, 132 kDa; ribonuclease A: 14 kDa; see inset). (c) Hydrodynamic radius distribution for PRORP1 (green) and PRORP2 (blue) in dynamic light scattering, confirming the monodispersity of PRORP samples. (d) SRCD analysis of PRORP1 (green) and PRORP2 (blue). SRCD spectra show a dominant peak at 190–200 nm characteristic of α -helices. The evaluation of two-dimensional structure content indicates 36/39% of α -helices, 15/16% of β -strands in PRORP1/PRORP2, respectively.

that this region is arranged in a super-helix as described in structurally related TPR proteins²⁷, which give the highest score in structure prediction with Phyre2. The C-terminal NYN domain could be modelled based on the structure of the MCPIP1 RNase that adopts an α - β PIN-like/NYN architecture. The Asp residues that are conserved in PRORP sequences are essential to the activity of MCPIP1 (see Supplementary Fig. S2) and their mutation abolished PRORP activity^{9,28}. This observation validates the proposed fold. The region that connects the N- and C-terminal domains was identified as a potential zinc-binding motif. Our mutational analysis confirmed that the two conserved Cys residues are involved in metal binding, together with another Cys and a His residue at the C-terminal end of PRORP. Overall, this results in a compact two-domain enzyme, as confirmed by the SAXS analysis in solution, with a zinc ion bridging the central and the C-terminal region (Fig. 5). Very recently, Howard and colleagues published a crystal structure of PRORP1 from *A. thaliana*²⁹. It confirms our structural predictions, in particular the superhelical fold of the PPR domain made of 5–6 PPR and PPR-like elements. SAXS data collected in solution on PRORP1 and PRORP2 show a good agreement (experimental and theoretical curves fit with χ of 4.9 and 2.8, respectively, for data in the range $0.02 < q < 0.2 \text{ \AA}^{-1}$)

with the atomic model (PDB id: 4g26). This validates the overall PRORP architecture with two functional domains: a N-terminal RNA-binding PPR domain and a C-terminal PIN-like catalytic domain, bridged together by a bipartite zinc-binding module. The four residues identified by mutagenesis and inductively coupled plasma mass spectrometry as zinc binders are also confirmed by the crystal structure.

Both footprint data and activity assays indicated that the tRNA precursor is essentially recognized by its acceptor arm, whereas the anticodon domain is dispensable. Results suggest that PRORP substrate recognition might be mediated by a limited number of determinants. As PRORP is able to recognize any tRNA of canonical structure, these determinants should be found among highly conserved residues such as G18 in loop D, C56 and R57 in loop T ψ C³⁰, which is corroborated by our results. Considering the length of the acceptor arm (45 Å) and the estimation of PRORP dimensions in SAXS (30 Å × 70 Å × 110 Å), the PPR domain is very likely to interact specifically with the D-T ψ C region at the corner of the tRNA, while the NYN catalytic domain must be located in the vicinity of the 5'-cleavage point. Thus, the proposed model (Fig. 5) shows the two-domains of PRORP2 that sandwich the substrate, their respective position acting as a ruler to determine the correct position of maturation, independently

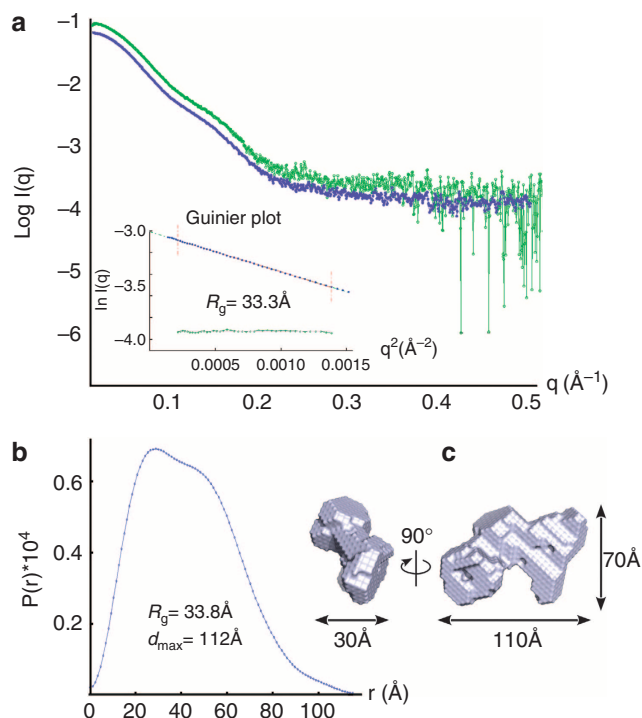


Figure 4 | SAXS analysis (a) PRORP1 and PRORP2 produce very similar intensity curves shown in green and blue, respectively. The inset Guinier plot indicates a gyration radius of $33.3 \pm 0.1 \text{ \AA}$. (b) This value is confirmed by the distance distribution function $P(r)$, which also suggests that PRORP proteins are composed of two distinct domains and an extended tail ($d_{\text{max}} > 100$). (c) Molecular envelope of PRORP2 derived from SAXS data analysis.

Table 1 | Inductively coupled plasma mass spectrometry identifies zinc in association with PRORP.

PRORP1	Zn ⁶⁶ (p.p.b.)	2 σ	Stoichiometry
WT	29.49	1.53	1
C344A	23.82	0.46	0.8
C347A	20.89	0.24	0.7
H548A	11.80	0.50	0.4
C565A	7.44	0.21	0.2
DD474-475AA	28.52	1.49	1
Buffer	0.44	0.12	0

Measurements were performed on wild-type PRORP1 (WT), as well as on proteins with point mutations applied to positions predicted to form the zinc-binding pocket. 2 σ indicates the s.e. in four replicate measurements.

from the internal sequence of the acceptor arm. Our data suggests an intriguing similarity in the mode of binding of the tRNA with the RNP RNase P. Indeed, earlier work has shown that bacterial RNase P interaction with the D-T ψ C region influences substrate binding and cleavage³¹. In the same line and similar to PRORP, bacterial and human RNP RNase P did not require the anticodon domain of tRNA for substrate recognition³². However, *E. coli* RNase P, contrary to PRORP, still allowed RNase P activity on a tRNA lacking its D domain³². The mechanistic model of the novel protein-only RNase P represents a good basis for further investigations of PRORP mode of action by complementary approaches, the ultimate step being the determination of a crystal structure of an active complex of PRORP and tRNA at atomic resolution.

The concept of structural mimicry of nucleic acids by proteins is well established, it has already been observed over 15 years ago^{33–35}. The specific case of PRORP is particularly interesting because both a single eukaryotic protein and a considerably more ancient bacterial ribozyme share the same catalytic function and appear to share similar RNA recognition processes. This implies that PRORP could represent an example of convergent evolution, with proteins that have evolved a mechanism of RNA recognition similar to that of catalytic RNA. This opens appealing perspectives for our understanding of the transition between the envisaged pre-biotic RNA world and the modern world dominated by proteins.

Methods

PRORP purification and characterization. *Arabidopsis* recombinant PRORP1 and PRORP2 proteins were expressed in *E. coli* and purified to homogeneity using affinity chromatography as described previously⁹. Before biophysical analyses (see below), a second step of size exclusion chromatography using a Superdex 200 10/300 GL column (GE Healthcare) was introduced to improve the quality of PRORP enzymes and to elute them in appropriate buffers. Proteins were concentrated by ultrafiltration to about 10 mg ml^{-1} , ultracentrifuged and stored at $4 \text{ }^\circ\text{C}$ until use in 50 mM HEPES-Na pH 7.5, 250 mM NaCl, 15% glycerol (w/v), 1 mM tris(2-carboxyethyl)phosphine. Sample homogeneity and particle size were systematically verified using dynamic light scattering (Malvern Zetasizer) at $20 \text{ }^\circ\text{C}$. Mass determination was performed by multi-angle light scattering using a SEC Superdex 200 column coupled to a Treos instrument (Wyatt technologies) in the storage buffer with 2% glycerol (w/v) only.

RNase P activity assays. cDNAs representing variants of *Arabidopsis* mitochondrial tRNA^{Cys} precursors were designed with leader and trailer sequences of 50 and 30 nucleotides, respectively, cloned in pUC19, transcribed *in vitro* by T7 RNA polymerase. tRNA^{Cys} precursor mutants included tRNAs with the anticodon domain removed (ΔAC), without both anticodon and D domain (ΔDAC), with point mutations at position 1, 18, 19, 56, 57 and 72. Sequences of oligonucleotide used to generate these mutants are available in Supplementary Table. For RNase P cleavage assays, reactions were always performed with three replicates using $0.5 \text{ }\mu\text{M}$ transcript and $0.15 \text{ }\mu\text{M}$ protein for 15 min at $25 \text{ }^\circ\text{C}$ as previously described⁹. RNA fragments were separated by denaturing PAGE and visualized by ethidium bromide staining. Quantifications were performed as described¹⁰.

Footprinting analyses. Recombinant PRORP1 was put in presence of equimolar amounts of 5'-³²P- γ -ATP radiolabeled mitochondrial tRNA^{Cys} precursors to form a PRORP/tRNA complex. As PRORP and tRNA only interact in a transient manner, the complex obtained was UV-crosslinked for 15 min at 260 nm. Samples were submitted to partial RNase V1 ($0.1 \text{ U } \mu\text{l}^{-1}$), RNase T1 ($1 \text{ U } \mu\text{l}^{-1}$) and RNase A ($1 \text{ }\mu\text{g } \mu\text{l}^{-1}$) digestions in the presence of competitor yeast RNA according to the manufacturer's instructions (Ambion, USA). The radiolabeled tRNA probe was also subjected to partial RNase T1 digestion in denaturing condition and to partial alkaline hydrolysis to generate RNA ladders. RNA samples were recovered by phenol/chloroform extractions, separated by high resolution 8% denaturing polyacrylamide gel electrophoresis and signal was acquired with a FLA-7000 phosphorimager (Fujifilm).

SRCD analysis. SRCD experiments were performed on the DISCO beamline at synchrotron SOLEIL (Saint-Aubin, France). The instrument was calibrated for magnitude and polarization with a 6.1-mg ml^{-1} D-10-camporsulfonic acid solution. PRORP proteins (10 mg ml^{-1}) in 100 mM potassium phosphate, 50 mM KCl, 10% (v/v) glycerol and 1 mM tris(2-carboxyethyl)phosphine were placed in a SRCD CaF₂ cuvette of $8 \text{ }\mu\text{m}$ pathlength. Three spectra between 170 and 280 nm were measured at 10, 20, 30, 40, 50, 60 and $70 \text{ }^\circ\text{C}$ to assess the thermal stability of PRORP1 and PRORP2. Data were processed (spectrum averaging, solvent base line subtraction) using CDtools³⁶. The secondary structure content of PRORPs was evaluated using the VARSLC method in DICHROWEB³⁷.

Small-angle X-ray scattering analysis. Small-angle X-ray scattering (SAXS) experiments were conducted on the SWING beamline at Synchrotron SOLEIL, Saint-Aubin, France. The beam wavelength was set to $\lambda = 1.033 \text{ \AA}$. The $17 \times 17 \text{ cm}^2$ low-noise Avici CCD detector was positioned at a distance of 2107 mm from the sample, with the direct beam off-centred. The resulting exploitable q -range was $0.005\text{--}0.5 \text{ \AA}^{-1}$, where $q = 4\pi \sin \theta / \lambda$, and 2θ is the scattering angle. PRORP samples at 10 mg ml^{-1} in 100 mM HEPES-Na (pH 7.5), 250 mM NaCl, 5% glycerol and 1 mM tris(2-carboxyethyl)phosphine were analysed by direct injection or high-performance liquid chromatography mode. In the first case, they were transferred into the SAXS flow-through capillary cell and a series of 50 frames was recorded. In the second case, they were loaded into a size exclusion column (Agilent Bio SEC-3,

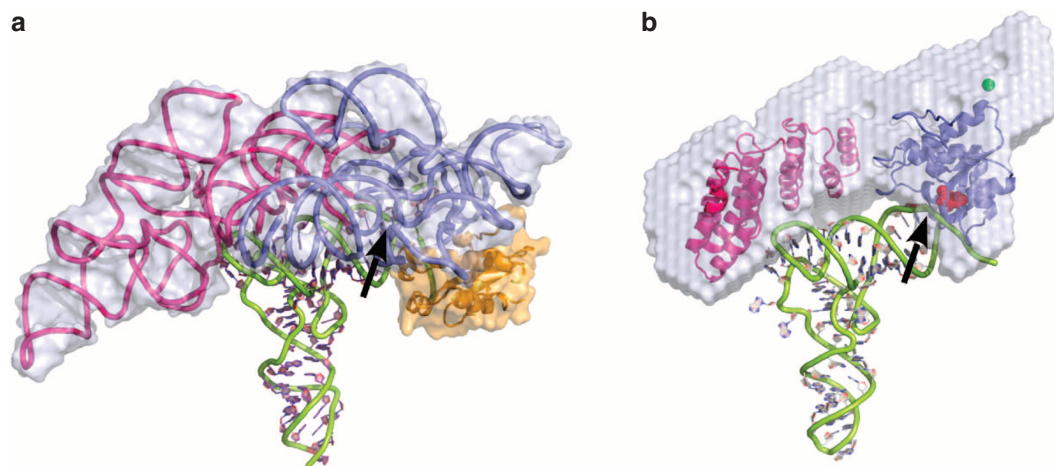


Figure 5 | A common mode of RNA binding in RNP and PRORP RNase P. (a) Structure of *Thermotoga maritima* ribozyme (PDBid 3Q1R¹³) with the catalytic domain in blue, the specificity domain in violet, the RNase P protein subunit in orange, the tRNA product in green and the molecular surface of the RNP in grey. (b) Model of *A. thaliana* PRORP2 protein shown in the same orientation and same colour code: the catalytic domain (NYN domain) in blue, with conserved Asp residues shown as red spheres adjacent to the 5'-cleavage site (indicated by an arrow), and the RNA-binding domain (PPR domain) in violet, interacting with the region of the D-TψC loops, were positioned in the SAXS envelope. The proposed zinc site is indicated by a green sphere in the central region, which connects the two main PRORP domains. This two-domains architecture offers a concave surface, which can be docked on the tRNA acceptor arm. Our data indicate that PRORP proteins have evolved an RNA recognition process very similar to that of RNP RNase P.

300 Å, 4.6 × 300 mm, 3 μm) using an Agilent high-performance liquid chromatography system and eluted into the SAXS flow-through capillary cell at a flow rate of 0.2 ml min⁻¹. SAXS measurements were collected throughout the whole protein elution time, with a frame duration of 1000 ms and a dead time between frames of 500 ms. Data processing, analysis and modelling steps were carried out with PRIMUS³⁸, and other programs of the ATSAS suite³⁹. The radius of gyration R_g was derived from Guinier approximation⁴⁰ and calculated from entire scattering pattern using the indirect transform package GNOM⁴¹, which provides the distance distribution function $P(r)$ of the particle. Based on this distribution, *ab initio* modelling was carried out with DAMMIF³⁹. A series of 11 dummy atom models was generated that were compared using the DAMAVER suite⁴² to determine the most typical/probable one (that is, showing the lowest averaged normalized spatial discrepancy). The molecular envelope corresponding to this model was used to spatially restrain the positions of PRORP domains and of the tRNA substrate in the model.

Inductively coupled plasma mass spectrometry. For the (w/v) analysis of metal cofactors, PRORP solutions resuspended in 0.5 N nitric acid were analysed with a ThermoElectron X Series II inductively coupled plasma mass spectrometry mass spectrometer operated at 1450 W, with argon carrier gas flow rate of 0.85 l min⁻¹, argon auxiliary gas flow rate of 0.40 l min⁻¹, using a Meinhardt quartz nebulizer, a quartz spray chamber with impact bead chilled to 3 °C and sample flow rate set to 0.1 l min⁻¹. Four replicate measurements were performed and values were corrected by an internal¹¹⁵In standard.

Structure modelling. The overall architecture of PRORP domains was predicted by homology modelling based on the alignment of 181 PRORP ortholog sequences⁹ and fold recognition to find remotely related candidates with known structure as implemented on the Phyre2 server²⁴. PRORP RNA partner (pre-tRNA^{Cys} from *A. thaliana*) was modelled using S2S⁴³ based on a sequence alignment with the tRNA^{Cys} from *E. coli* (PDB-id 1B23⁴⁴). PRORP domain models were fit in the SAXS envelope and the tRNA substrate was docked on its concave surface in a way bringing the PPR and catalytic domains in close vicinity of the D-TψC corner and of the cleavage point, respectively. Molecular docking and related figures were performed with PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0, Schrödinger, LLC).

References

- Hartmann, E. & Hartmann, R. K. The enigma of ribonuclease P evolution. *Trends Genet.* **19**, 561–569 (2003).
- Lai, L. B., Vioque, A., Kirsebom, L. A. & Gopalan, V. Unexpected diversity of RNase P, an ancient tRNA processing enzyme: challenges and prospects. *FEBS Lett.* **584**, 287–296 (2010).
- Jarrous, N. & Gopalan, V. Archaeal/eukaryal RNase P: subunits, functions and RNA diversification. *Nucleic Acids Res.* **38**, 7885–7894 (2010).
- Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**, 849–857 (1983).
- Altman, S. A view of RNase P. *Mol. Biosyst.* **3**, 604–607 (2007).
- Wang, M. J., Davis, N. W. & Gegenheimer, P. Novel mechanisms for maturation of chloroplast transfer RNA precursors. *EMBO J.* **7**, 1567–1574 (1988).
- Rossmannith, W. & Karwan, R. M. Characterization of human mitochondrial RNase P: novel aspects in tRNA processing. *Biochem. Biophys. Res. Commun.* **247**, 234–241 (1998).
- Holzmann, J. *et al.* RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell* **135**, 462–474 (2008).
- Gobert, A. *et al.* A single Arabidopsis organellar protein has RNase P activity. *Nature Struct. & Molec. Biol.* **17**, 740–744 (2010).
- Gutmann, B., Gobert, A. & Giegé, P. PRORP proteins support RNase P activity in both organelles and the nucleus in Arabidopsis. *Genes Dev.* **26**, 1022–1027 (2012).
- Taschner, A. *et al.* Nuclear RNase P of *Trypanosoma brucei*: A single protein in place of the multicomponent RNA-protein complex. *Cell Reports* **2**, 19–25 (2012).
- Westhof, E. & Altman, S. Three-dimensional working model of M1 RNA, the catalytic RNA subunit of ribonuclease P from *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **91**, 5133–5137 (1994).
- Reiter, N. J. *et al.* Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature* **468**, 784–789 (2010).
- Schmitz-Linneweber, C. & Small, I. Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends Plant Sci.* **13**, 663–670 (2008).
- Anantharaman, V. & Aravind, L. The NYN domains: novel predicted RNases with a PIN domain-like fold. *RNA Biol.* **3**, 18–27 (2006).
- Thomas, B. C., Li, X. & Gegenheimer, P. Chloroplast ribonuclease P does not utilize the ribozyme-type pre-tRNA cleavage mechanism. *RNA* **6**, 545–553 (2000).
- Pavlova, L. V. *et al.* tRNA Processing by protein-only versus RNA-based RNase P: kinetic analysis reveals mechanistic differences. *ChemBiochem.* (2012).
- Giegé, R. *et al.* Structure of transfer RNAs: similarity and variability. *Wiley Interdiscip. Rev. RNA* **3**, 37–61 (2012).
- Putnam, C. D., Hammel, M., Hura, G. L. & Tainer, J. A. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **40**, 191–285 (2007).
- Lipfert, J. & Doniach, S. Small-angle X-ray scattering from RNA, proteins, and protein complexes. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 307–327 (2007).
- Mertens, H. D. & Svergun, D. I. Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J. Struct. Biol.* **172**, 128–141 (2010).
- Zwieb, C. *et al.* Structural modeling of RNase P RNA of the hyperthermophilic archaeon *Pyrococcus horikoshii* OT3. *Biochem. Biophys. Res. Commun.* **414**, 517–522 (2011).

23. Kazantsev, A. V. *et al.* Solution structure of RNase P RNA. *RNA* **17**, 1159–1171 (2011).
24. Kelley, L. A. & Sternberg, M. J. E. Protein structure prediction on the web: a case study using the Phyre server. *Nat. Protoc.* **4**, 363–371 (2009).
25. Lurin, C. *et al.* Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* **16**, 2089–2103 (2004).
26. Small, I. D. & Peeters, N. The PPR motif—a TPR-related motif prevalent in plant organellar proteins. *Trends Biochem. Sci.* **25**, 46–47 (2000).
27. Blatch, G. L. & Lassle, M. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays* **21**, 932–939 (1999).
28. Xu, J. *et al.* Structural study of MCPIP1 N-terminal conserved domain reveals a PIN-like RNase. *Nucleic Acids Res.* **40**, 6957–6965 (2012).
29. Howard, M. J., Lim, W. H., Fierke, C. A. & Koutmos, M. Mitochondrial ribonuclease P structure provides insight into the evolution of catalytic strategies for precursor-tRNA 5' processing. *Proc. Natl Acad. Sci. USA* **109**, 16149–16154 (2012).
30. Westhof, E. & Auffinger, P. *Transfer RNA structure eLS* (John Wiley & Sons, Ltd, 2012).
31. Brannvall, M., Kikovska, E., Wu, S. & Kirsebom, L. A. Evidence for induced fit in bacterial RNase P RNA-mediated cleavage. *J. Mol. Biol.* **372**, 1149–1164 (2007).
32. Hardt, W. D., Schlegl, J., Erdmann, V. A. & Hartmann, R. K. Role of the D arm and the anticodon arm in tRNA recognition by eubacterial and eukaryotic RNase P enzymes. *Biochemistry* **32**, 13046–13053 (1993).
33. Putnam, C. D. *et al.* Protein mimicry of DNA from crystal structures of the uracil-DNA glycosylase inhibitor protein and its complex with *Escherichia coli* uracil-DNA glycosylase. *J. Mol. Biol.* **287**, 331–346 (1999).
34. Mol, C. D. *et al.* Crystal structure of human uracil-DNA glycosylase in complex with a protein inhibitor: protein mimicry of DNA. *Cell* **82**, 701–708 (1995).
35. Nissen, P., Kjeldgaard, M. & Nyborg, J. Macromolecular mimicry. *EMBO J.* **19**, 489–495 (2000).
36. Lees, J. G., Smith, B. R., Wien, F., Miles, A. J. & Wallace, B. A. CDtool—an integrated software package for circular dichroism spectroscopic data processing, analysis, and archiving. *Anal. Biochem.* **332**, 285–289 (2004).
37. Whitmore, L. & Wallace, B. A. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* **32** (2004).
38. Konarev, P., Volkov, V., Sokolova, A., Koch, M. & Svergun, D. PRIMUS: a windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277–1282 (2003).
39. Konarev, P., Petoukhov, M., Volkov, V. & Svergun, D. ATSAS 2.1: a program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **39**, 277–286 (2006).
40. Guinier, A. La diffraction des rayons X aux très petits angles: application à l'étude de phénomènes ultramicroscopiques. *Ann. Phys. (Paris)* **12**, 161–237 (1939).
41. Svergun, D. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Cryst.* **25**, 495–503 (1992).
42. Volkov, V. V. & Svergun, D. I. Uniqueness of *ab-initio* shape determination in small-angle scattering. *J. Appl. Cryst.* **36**, 860–864 (2003).
43. Jossinet, F. & Westhof, E. Sequence to structure (S2S): display, manipulate and interconnect RNA data from sequence to structure. *Bioinformatics* **21**, 3320–3321 (2005).
44. Nissen, P., Thirup, S., Kjeldgaard, M. & Nyborg, J. The crystal structure of Cys-tRNA(Cys)-EF-Tu-GDPNP reveals general and specific features in the ternary complex and in tRNA. *Struct. Fold Des.* **7**, 143–156 (1999).

Acknowledgements

We thank Bernard Lorber for critical reading of the manuscript and assistance in dynamic light scattering experiments, as well as Isabelle Billas Massobrio for multi-angle light scattering analysis. We thank Elodie Ubrig and Laurie-Anne Roeckel for technical assistance and Pierre Fechter for advice with footprinting experiments. We also acknowledge the staff of SOLEIL synchrotron (Saint-Aubin, France) for the beamtime allocated to the project, and more particularly Frank Wien for assistance during SRCD data collection on DISCO. This work was supported by the French 'Centre National de la Recherche Scientifique' and by the University of Strasbourg. AG and OF were supported by an ANR Blanc research grant 'PRO-RNase P, ANR 11 BSV8 008 01' to PG and CS and by the LabEx consortium 'MitoCross'. FP and BG were supported by PhD grants from the University of Strasbourg.

Author contributions

A.G., C.S. and P.G. designed and coordinated the experiments. CS directed the biophysical characterization of PRORP and PG directed the biochemical study of the PRORP/tRNA complex. A.G., F.P., O.F., B.G., R.B., P.R., C.S. and P.G. performed the experiments and analysed the results. A.G., F.P., C.S. and P.G. wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Gobert, A. *et al.* Structural insights into protein-only RNase P complexed with tRNA. *Nat. Commun.* 4:1353 doi: 10.1038/ncomms2358 (2013).



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>