Structural and mechanistic basis of RNA processing by protein-only ribonuclease P enzymes

Arjun Bhatta1,2 and Hauke S. Hillen1,2,3,*,@  

Ribonuclease P (RNase P) enzymes are responsible for the 5′ processing of tRNA precursors. In addition to the well-characterised ribozyme-based RNase P enzymes, an evolutionarily distinct group of protein-only RNase Ps exists. These proteinaceous RNase Ps (PRORPs) can be found in all three domains of life and can be divided into two structurally different types: eukaryotic and prokaryotic. Recent structural studies on members of both families reveal a surprising diversity of molecular architectures, but also highlight conceptual and mechanistic similarities. Here, we provide a comparison between the different types of PRORP enzymes and review how the combination of structural, biochemical, and biophysical studies has led to a molecular picture of protein-mediated tRNA processing.

Protein-based RNase Ps in all domains of life  

tRNAs mediate the translation of genetic information from nucleotide sequences into protein molecules, and thus play a fundamental role in gene expression [1–3]. They are transcribed as precursors from their respective genes, and must undergo several maturation steps before they are functionally competent [4–6]. In almost all organisms, this entails the endonucleolytic removal of a leader sequence at the 5′ end of the tRNA. This processing step is catalysed by a class of enzymes called RNase P [6–8].

The first RNase P enzymes discovered were ribozymes, consisting of a catalytic RNA and a varying number of accessory proteins [3–13]. Such ribonucleoprotein-based RNase Ps (RNP RNase Ps; see Glossary) are found in all three domains of life and were long presumed to be the only form of RNase P [12,14]. This was challenged in 1988, when Peter Gegenheimer and colleagues showed that plant chloroplast RNase P exhibits protein-like behaviour and is insensitive to micrococcal nuclease, thus suggesting a protein-based RNase P enzyme devoid of any RNA component [15,16]. Subsequent characterisations of RNase P preparations from mammalian and trypanosomatid mitochondria suggested similar protein-catalysed RNase P activity in these organelles [17–19].

These observations were initially met with scepticism [20], and it was not until 2008 that the existence of an entirely PRORP was unequivocally demonstrated by Walter Rossmanith and colleagues. His group showed that human mitochondrial RNase P (mtRNase P) is a multisubunit protein complex, identified its constituent subunits, and demonstrated protein-catalysed RNase P activity of this complex both in vitro and in vivo [21]. Homologues of the endonuclease subunit of mtRNase P were subsequently identified in four out of five eukaryotic supergroups [22]. In contrast to mtRNase P, these homologues outside the metazoa lineages appear to be single-subunit enzymes that are catalytically active without the requirement of additional proteins [23,24]. In addition to the eukaryotic PRORP enzymes, a second form of PRORP was recently discovered in Aquifex aeolicus.
(aQRNase P), a bacterium that possesses neither an RNP RNase P nor a eukaryotic-type protein-only RNase P [25]. Instead, it uses a minimal prokaryotic form of protein-only RNase P that is comprised only of a single metalloenzyme domain. Homologues of A. aeolicus RNase P (HARPs) were also identified in some bacteria and many archaea [25], indicating that this type of protein-only RNase P is widespread among the prokaryotic and archaeal lineages. Taken together, these discoveries establish the existence of two distinct archetypes of protein-only RNase Ps, eukaryotic PRORPs and prokaryotic HARPs, spanning across all domains of life.

The mechanism of ribozyme-catalysed tRNA 5′-processing has been elucidated in detail through structural and functional studies [26–38]. By contrast, the molecular basis of RNA processing by protein-only RNase P enzymes long remained speculative. Over the past two decades, a large body of biochemical and biophysical studies have provided insights into the mechanisms of substrate recognition and catalysis by protein-only RNase P enzymes. Recently, high-resolution structures of members of each of the different types of protein-only RNase Ps have been reported (Table 1). From the combination of functional and structural studies, a molecular picture of tRNA processing by PRORPs has begun to emerge. In this review, we summarise our current understanding of protein-catalysed tRNA 5′-processing and highlight both similarities and differences between the different types of protein-only RNase P enzymes.

Molecular architectures of protein-only RNase P enzymes

Eukaryotic protein-only RNase Ps

Eukaryotic PRORPs all share a characteristic tripartite domain architecture comprised of an N-terminal pentatricopeptide repeat (PPR) domain, a split zinc-binding domain (ZBD), and a C-terminal PILT N terminus (PIN)-like nuclease domain [22,39,40]. While most PRORPs are single-subunit enzymes (ssPRORPs), the PRORP enzymes in metazoan mitochondria are multisubunit protein complexes (msPRORPs).

Single-subunit protein-only RNase Ps

The most well-studied ssPRORPs are those of Arabidopsis thaliana and Trypanosoma brucei. Both organisms do not encode an RNP RNase P and instead rely exclusively on ssPRORP enzymes for 5′-processing [22–24,41]. Nuclear and organellar tRNA processing are carried out by a structural element of tRNA formed by base pairing of T and D loops of tRNA. Hydrogen bonding interactions between these two elements are required for catalytic activity. Hydroxylating enzymes catalysing the transfer of a methyl (-CH3) group from a cofactor to a substrate. Pentatricopeptide repeat (PPR): a 35-amino-acid long motif, adopting a helix-turn-helix structure. Typically, multiple PPR repeats are present in tandem and are involved in RNA binding. PPR code: molecular principle that has been proposed for the modular, sequence-specific recognition of RNA bases by plant PPR proteins [67,68]. Ribonuclease (RNPS): a macromolecular complex comprised of RNA and protein subunits. T arm: a stem-loop structure formed by the nucleotides at positions 49–65 in a tRNA structure. D arm: a stem-loop structure formed by the nucleotides at positions 1–7 and 66–72 in a tRNA structure.

A. thaliana PRORP1 PPR domain in complex with tRNA

Yeast tRNA

6LVR

Table 1. Structures of protein-only RNase P enzymes

<table>
<thead>
<tr>
<th>Protein or complex</th>
<th>RNA</th>
<th>PDB code</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabidopsis thaliana PRORP1</td>
<td>–</td>
<td>4G24</td>
<td>[39]</td>
</tr>
<tr>
<td>A. thaliana PRORP2</td>
<td>–</td>
<td>5DIZ</td>
<td>[42]</td>
</tr>
<tr>
<td>A. thaliana PRORP2</td>
<td>–</td>
<td>5FT9</td>
<td>[43]</td>
</tr>
<tr>
<td>Homo sapiens PRORP - free</td>
<td>–</td>
<td>4ROU</td>
<td>[57]</td>
</tr>
<tr>
<td>H. sapiens PRORP - free</td>
<td>–</td>
<td>4XGL</td>
<td>[53]</td>
</tr>
</tbody>
</table>
| A. thaliana PRORP1 PPR domain in complex with tRNA     | Yeast tRNA

6LVR

[43] |
| Aquifex aeolicus RNase P                                | –     | 7F3E     | [61] |
| Halorhodospira halophila RNase P                        | –     | 7OG5     | [62] |
| Thermococcus celer RNase P                              | –     | 7E8U     | [66] |
| Planctomyces bacterium RNase P                          | –     | 7E9K     | [65] |
| Planctomyces bacterium RNase P in complex with tRNA    | Escherichia coli tRNA

7E8O

[66] |
| H. sapiens mtRNase P in complex with tRNA               | H. sapiens mt-tRNA

7ONU

[59] |
by distinct isoforms of ssPRORP, which share a conserved domain organisation and up to 30% sequence similarity between the two organisms and with human PRORP [23,24,41]. A. thaliana and T. brucei PRORPs are catalytically active without the requirement of any additional subunits in vitro and in vivo, and are therefore bona fide ssPRORPs [23,24,41].

The molecular architecture of ssPRORPs has been elucidated through crystal structures of two A. thaliana PRORP isoforms, PRORP1 and PRORP2 (atPRORP1 and atPRORP2) (Figure 1A,B) [39,42,43]. These structures reveal a V-shaped arrangement of the three domains, in which the PPR and nuclease domains form the two arms of the ‘V’, respectively, and the ZBD forms the central vertex. The nuclease domain adopts a PIN/NYN domain-like fold with structural resemblance to the nuclease domains of DNA polymerase I and FLAP nucleases [39,40,44,45]. Its active site contains four catalytic aspartate residues that are highly conserved among all PRORPs [22,39,42]. The ZBD is comprised of a four-stranded antiparallel β sheet, which contains three invariant cysteines and one histidine involved in coordination of a Zn²⁺ ion [39]. The PPR domain is composed of 11 α helices, which together constitute five complete PPR repeats as well as one incomplete repeat. These repeats form a curved structure with an inner concave surface that faces the active site of the nuclease domain. Since many PPR proteins are RNA-binding proteins, the PPR domain was proposed to be involved in substrate binding [39,43,46–48]. This was recently confirmed by a structure of the atPRORP1 PPR domain in complex with tRNA, which shows that the PPR domain forms specific contacts with the RNA [49]. Comparison with the previous substrate-free structures indicates that the V-shaped arrangement of PPR and nuclease domains would need to widen to

Figure 1. Structures of eukaryotic protein-only ribonuclease (RNase) Ps.

(A) Structure of Arabidopsis thaliana proteinaceous RNase P PRORP1 (atPRORP1) (PDB: 4G24) [39]. A schematic domain representation is shown above the structure with individual domains coloured in shades of green. Protein regions modelled in the structure are indicated by a black line above the schematic.

(B) Structure of A. thaliana PRORP2 (atPRORP2) (PDB: 5DIZ) [42]. Domain representation and colouring as in (A).

(C) Structure of substrate-bound human mitochondrial RNase P complex (PDB: 7ONU) [59]. A schematic domain representation of the nuclease subunit Homo sapiens proteinaceous RNase P (hsPRORP) is shown above the structure with the same colouring as in (A). The additional subunits TRMT10C and SDR5C1 are coloured in shades of blue and grey, respectively. The tRNA precursor is shown in red. The longer extensions in atPRORP1 and hsPRORP1 in (A) and (C) correspond to their respective organelar targeting sequences. Abbreviation: PPR, pentatricopeptide repeat.
accommodate the tRNA between the two domains [49]. This is supported by small-angle X-ray scattering (SAXS) studies of full-length atPRORP in complex with tRNA and in silico normal mode analysis, which suggest that atPRORPs possess conformational flexibility [42,43,48]. Thus, ssPRORPs may undergo rearrangements during substrate binding.

In summary, the structures of atPRORP enzymes reveal the conserved molecular architecture of ssPRORPs and provide insights into their mechanism of substrate binding.

Metazoan multisubunit protein-only RNase P
In contrast to the widespread ssPRORPs, msPRORPs have thus far only been found in mitochondria of metazoans [21,50]. In humans, the mtRNase P is a trimeric complex comprised of an endonuclease, a methyltransferase, and a fatty acid/sterol dehydrogenase [21,51,52]. Its nucleolytic subunit Homo sapiens PRORP (hsPRORP; also known as MRPP3) is an orthologue of ssPRORPs and exhibits a high degree of sequence similarity to the ssPRORPs from A. thaliana and T. brucei [23,53]. In contrast to these, however, hsPRORP requires the two additional mtRNase P subunits, TRMT10C and SDR5C1, for tRNA processing [21]. TRMT10C (tRNA methyltransferase 10 type C; also known as MRPP1) is a SPOUT-fold methyltransferase which is responsible for methylating mitochondrial tRNAs, while SDR5C1 (short-chain dehydrogenase/reductase family 5C, member 1; also known as MRPP2 or HSD17B10) is a dehydrogenase involved in β oxidation of a diverse array of substrates [51,52,54]. TRMT10C and SDR5C1 interact with each other independent of hsPRORP to form a complex that can bind and methylate the pre-tRNA [52,55,56].

The first structural insights into human mtRNase P came from two crystal structures of hsPRORP, which reveal a V-shaped organisation resembling ssPRORPs (Figure 2A) [53,57]. In both structures, the active site of the nuclease domain is partially disordered, and catalytic residues are engaged in intra-molecular interactions. While these interactions differ between the two structures, the authors of both studies concluded that hsPRORP adopts an autoinhibited conformation in isolation. However, both studies used truncated variants of hsPRORP lacking parts of the PPR domain that are inactive in biochemical assays even in the presence of TRMT10C and SDR5C1 [53,57], which has raised concerns about the physiological relevance of the observed states.
Whether free full-length hsPRORP also adopts an autoinhibited conformation remains to be determined experimentally. Taken together, these structures confirmed hsPRORPs’ similarity to plant ssPRORPs, and suggested that it may adopt a catalytically incompetent conformation prior to formation of the substrate-engaged mtRNase P complex.

Recently, the cryoelectron microscopy (cryo-EM) structure of the complete human mtRNase P complex bound to pre-tRNA was reported (Figure 1C) [59]. This structure reveals how hsPRORP, TRMT10C, and SDR5C1 interact with each other and with the pre-tRNA substrate. SDR5C1 forms a homo-tetramer, which interacts with TRMT10C to form a platform to which the tRNA substrate binds through multiple interactions. This positions the methyl acceptor, a purine base in position 9 of the tRNA, in the methyltransferase domain of TRMT10C. HsPRORP binds atop this complex, with the scissile phosphodiester bond at the 5′ end of the tRNA positioned in the nuclease active site. Compared with the previous structures, hsPRORP adopts a more open conformation in which the autoinhibitory interactions are abolished and the active site organisation resembles that observed in arPRORP1 [59]. This conformation appears to be stabilised by interactions with TRMT10C and the tRNA, suggesting that hsPRORP may become activated upon recruitment to the complex (Figure 2B) [59]. The structure of mtRNase P thus reveals the active conformation of hsPRORP and shows how this multifunctional complex integrates two enzymatic activities into one molecular machinery.

Prokaryotic minimal protein-only RNase Ps

The most recent type of PRORP enzymes discovered are the prokaryotic protein-only RNase Ps, also known as HARPs. While HARPs appear to have functionally replaced RNP RNase Ps in some bacteria like A. aeolicus, they coexist alongside the RNP RNase Ps in many archaea [25,60]. HARPs represent the most minimal PRORP enzymes known to date, as they are comprised solely of a metalloenzyme domain that belongs to the same PIN domain-like superfamily as the nuclease domains of eukaryotic PRORPs [25,40,45].

Recent cryo-EM studies of HARPs from A. aeolicus and Halorhodospira halophila demonstrate that they adopt a different architecture than eukaryotic PRORPs [61,62]. They assemble into large dodecameric assemblies comprised of hexamers of dimers, wherein two monomers dimerise via their spike helix or protruding helix (SH/PrH) microdomains, and the resulting dimers interact side by side to form a superhelical assembly (Figure 3) [61,62]. The overall fold of the individual monomers is distinct from the nuclease domain of eukaryotic PRORPs, and instead shows similarities to different members of the PIN family, for example, the VapC4 toxin from Pyrococcus horikoshii [40,60,61,63,64]. This suggests that HARPs may have an independent evolutionary origin from eukaryotic PRORPs [61,62]. Biochemical experiments show that the oligomerisation of HARPs is necessary for their RNase P activity [62], and it has been proposed that two adjacent dimers may cooperate to mediate processing of a single tRNA molecule [61,62]. In this case, a tetrameric assembly formed by the association of two dimers would represent the minimal catalytic unit of HARPs [61,62]. This hypothesis is supported by a recently reported crystal structure of Planctomycetes bacterium HARPs (pHARPs) in complex with pre-tRNA [65]. As dodecameric assemblies, HARPs could contain up to ten substrate-binding sites, which may allow simultaneous processing of multiple tRNAs [61,62,65]. However, this model remains speculative, as a structure of substrate-engaged dodecameric HARP assembly has not yet been determined.

In summary, these structural studies have elucidated the molecular architecture of prototypic members of the different forms of protein-only RNase Ps. In conjunction with a large body of previous biochemical and biophysical data, this now enables a comparison of their molecular mechanisms of substrate recognition, specificity, and cleavage.
Substrate recognition and specificity

The PRORP enzymes generally interact with the same elements in tRNAs as RNP RNase Ps, but utilise distinct molecular binding principles \[48,49,59,61,62,65\]. Structures of RNP RNase Ps show that these enzymes recognise their substrates through specific interactions with the...
acceptor arm and the elbow of the tRNA, as well as the 5’ and 3’ extensions [29,32,34]. In particular, they recognise a conserved base pair (G/U19-C56) between the T arm and the D arm, which together form the elbow, through RNA–RNA stacking interactions (Figure 4A,B).

Substrate binding by single-subunit protein-only RNase Ps

The mechanism of substrate binding by ssPRORPs was elucidated through a combination of biochemical, biophysical, and structural studies. This revealed that the PPR domain interacts with the tRNA elbow in a conceptually similar manner as in RNP RNase Ps, mediated through two mechanisms (Figure 4C,D) [43,48,49]. First, it forms a positively charged binding pocket which accommodates the tRNA elbow and stabilises the phosphate backbone. Second, it forms base-specific interactions with the tRNA elbow. The structure of the atPRORP1 PPR domain in complex with yeast tRNA\textsuperscript{Phe} shows that residues Y133, Y140, and R210 interact with D/U17, the G18-Ψ/U55 base pair, and the G/U19-C56 base pair of the tRNA, respectively [49]. This is consistent with previous SAXS data and biochemical data, which show that substitution of these protein residues or bases leads to decreased substrate affinity and cleavage efficiency [43,47,48]. While no substrate-bound structures are available for other ssPRORPs, biochemical

Figure 4. Substrate recognition in ribonuclease (RNase) P enzymes. (A,C,E,G) Schematic representations of tRNAs from different organisms and cellular compartments, indicating the conservation of their secondary structure elements: (A) human nucleus, (C) Arabidopsis thaliana mitochondria, (E) Planctomycetes bacterium, and (G) human mitochondria [92,93]. Transparent grey circles represent individual nucleotides. Blue or purple circles indicate conserved nucleotides. G/U19-C56 interactions are coloured in purple. Acceptor arm, D arm, anticodon arm, variable region, and T arm are labelled Acc, D, AC, V, and T, respectively. Strong overlap and low transparency of circles indicates conservation, while weak overlap and high transparency indicates variability. Regions interacting with human nuclear ribonuclease P (tRNA\textsuperscript{Phe} P, A. thaliana proteinaceous RNase P (atPRORP1), Homo sapiens PRORP (hsPRORP), and hsTRMT10C are marked by brown, green, green, and blue arcs, respectively. (B) Interaction of human nuclear RNase P with tRNA (PDB: 6AHU) [32]. The catalytic RNA component is coloured in brown, and protein subunits are coloured in grey. Interactions with the tRNA elbow are shown in inset. Acceptor arm, D arm, anticodon arm, variable region, and T arm are coloured red, orange, maroon, light-brown, and magenta, respectively. Hydrogen bonds are shown as broken lines throughout. (D) Interaction of atPRORP1 pentatricopeptide repeat (PPR) domain with the tRNA elbow (PDB: 6LVR) [49]. The PPR domain is coloured in green. Interactions with the tRNA elbow are shown in inset. (F) Interaction of Planctomycetes bacterium RNase P with tRNA (PDB: 7E80) [65]. The nucleotide at position 56 is a G in the tRNA substrate used for this study, and G/U19-C56 interactions are not present. (H) Interaction of human mitochondrial RNase P (mtRNase P) complex with pre-tRNA (PDB: 7ONU) [59]. Colouring as in Figure 1C. Interactions with the tRNA elbow and the anticodon loop are shown in inset.
data indicate that atPRORP2 and atPRORP3 also form base-specific interactions between the tRNA elbow and the PPR domain [47,66]. Thus, although the precise interactions may differ, specific recognition of tRNA elbow bases appears to be a conserved feature among ssPRORPs. Notably, these interactions differ from the previously proposed PPR code derived from plant organellar PPR proteins, which bind RNAs in a one-base-per-repeat fashion [46,47,49,67–69]. Thus, PPR proteins appear to have diverged substantially with respect to their RNA binding mechanisms. Taken together, these observations suggest that ssPRORPs have adapted their PPR domain to recognise similar structural elements in tRNAs as the RNP RNase Ps.

Substrate binding by mitochondrial multisubunit protein-only RNase Ps

By contrast, the mechanism of substrate recognition by the msPRORPs in metazoan mitochondria appears to differ. The structure of human mtRNase P shows that the PPR domain of hsPRORP forms a similar charge-complementary binding groove for the tRNA elbow as in ssPRORPs, but it does not form base-specific interactions (Figure 4G, H) [59]. Instead, the additional subunits TRMT10C and SDR5C1 interact with the pre-tRNA by two different mechanisms. First, the TRMT10C/SDR5C1 complex forms a shape- and charge-complementary binding platform that interacts with all four arms of the tRNA via nonspecific backbone stabilisation. Second, TRMT10C additionally interacts with the anticodon loop through a specific interaction between its residue R181 and a conserved base in position 33 of the tRNA (Figure 4G, H) [59]. Thus, the additional subunits in metazoan mtRNase Ps enable a unique substrate recognition mechanism that involves interactions with tRNA elements that are not bound by any other known RNase P.

A possible explanation for the recognition of distinct elements by the mitochondrial msPRORPs may lie in the structural properties of their substrates. Metazoan mitochondrial tRNAs exhibit a large degree of variability in otherwise conserved tRNA elements [70–72] (Figure 4A, C, E, G). In particular, the sequences, lengths, and structures of their T and D arms are highly degenerate (Figure 4G). The canonical substrate recognition mechanism used by RNP RNase P and ssPRORP enzymes, which involves specific interactions with the elbow region, may thus not be feasible or sufficient for the binding of all metazoan mitochondrial tRNAs. By contrast, the base in the anticodon loop that interacts with TRMT10C is conserved as pyrimidine in mitochondrial tRNAs, and thus offers an alternative specific recognition element. This suggests that the additional subunits in metazoan mitochondrial msPRORPs may have been adopted to compensate for the structural variability in their tRNA substrates [8,21,59].

Substrate binding by prokaryotic minimal protein-only RNase Ps

Based on structural modelling, it was suggested that HARP may recognise similar tRNA elements as ssPRORPs [61,62]. In particular, it was proposed that the SH/PrH microdomain may interact with the elbow. This was recently confirmed by the crystal structure of pbHARP in complex with pre-tRNA, which shows that a positively charged patch formed by conserved arginine residues (R116, R123, R138, and R142 in pbHARP) interacts with the tRNA elbow via backbone interactions [65]. Mutation of these residues to alanine results in decreased activity of both pbHARP and hhHARP [62,65]. In contrast to ssPRORPs, however, no base-specific interactions between the tRNA elbow and HARP were observed (Figure 4E, F) [65]. It has been proposed that the binding of the elbow to the SH/PrH microdomain of one monomer of a dimer may position the 5’ end of the tRNA in the active site of a monomer in the adjacent dimer [61,62]. The distance between these two binding sites could then act as a molecular ruler to define the 5’ cleavage site, as has been analogously proposed for RNP RNase Ps and ssPRORPs [29,34,47]. However, this model remains speculative to date, as the 5’ cleavage site was not positioned in an active site in the crystal structure of pbHARP in complex with pre-tRNA [65]. Taken together, the available
data suggest that HARP sinteract with their substrates in a conceptually similar fashion as RNP RNase Ps and ssPRORPs, but may rely less on base-specific interactions with the tRNA elbow. However, further work is required to fully elucidate their molecular mechanism.

Substrate specificity of protein-only RNase P enzymes

The observed differences in substrate recognition strategies by the different forms of RNase Ps may also partially explain their substrate specificities. The multisubunit mtRNase P, which makes interactions with all four tRNA arms, cannot process mt-tRNA^{Ser(AGY)}, a mitochondrial tRNA lacking the entire D arm [73,74]. By contrast, many other RNase P enzymes can cleave non-tRNA substrates with structural similarities to tRNAs. For example, structured RNAs such as the 4.5S RNA and tmRNAs in *Escherichia coli*, or MALAT1 and MEN-β/NEAT1 in humans, have been described as substrates for RNP RNase P [75–77]. Several observations indicate that ssPRORPs and HARP s may also have a broad substrate spectrum. First, both are able to rescue the otherwise lethal knockout of endogenous RNP RNase P in yeast or *E. coli* [23,25,78,79]. Second, the ssPRORP isoform that localises to plant mitochondria has been shown to process t elements, noncanonical substrates encoded in the mitochondrial genome, both *in vivo* and *in vitro* [41]. Third, artificial substrates resembling tRNAs with truncated D arm or anticodon arm can be processed by *A. thaliana* PRORPs *in vitro* [47,80]. While processing of non-tRNA substrates by HARP s has not been demonstrated, their conceptually similar substrate binding mechanism suggests that they may have a similarly broad substrate spectrum as ssPRORPs.

Taken together, these structural and functional studies show that all types of protein-only RNase Ps interact with the tRNA elbow and the 5′ cleavage site. However, their detailed interaction strategies differ, and metazoan mtRNase Ps additionally interact with several other structural elements of tRNAs.

Mechanism of catalysis by protein-only RNase Ps

Despite their distinct structures and substrate recognition modes, all protein-based RNase Ps appear to use a common catalytic mechanism. In particular, they all share a conserved active site architecture. In *at*PRORP1 and *hs*PRORP, the active sites are comprised of four highly conserved aspartate residues, which coordinate catalytic metal ions near the cleavage site (Figure 5A,B) [39,59]. Three out of these four aspartates – D399, D475, and D493 in *at*PRORP1; D409, D479, and D499 in *hs*PRORP – are also strictly conserved among HARP s and adopt a similar organisation to coordinate two catalytic metal ions (Figure 5C) [61,62,65]. The fourth aspartate (D474 in *at*PRORP1, D478 in *hs*PRORP) is replaced by a glutamate residue in *hh*HARP and *aq*RNase P, but is not conserved in HARP s from other organisms, including *pb*HARP, and may thus not be essential for catalysis. Instead, HARP active sites contain a different conserved aspartate residue at position 138 in *aq*RNase P (D139 in *hh*HARP and D151 in *pb*HARP), which is absent in PRORPs. While these differences may reflect the distinct evolutionary origins of HARP s and PRORPs, the otherwise remarkable similarity in the active site organisation indicates conservation of the catalytic mechanism [25,61,62,65].

From biochemical and structural data on both single-subunit and multisubunit PRORPs, a mechanistic model for catalysis by protein-only RNase Ps has emerged [39,59,81–83]. In this model, the four active site aspartates coordinate two Mg^2+ ions, which in turn coordinate the pro-S₂ oxygen and the 3′ proximal phosphoester-bonded oxygen of the scissile phosphate in the RNA substrate. This is supported by the structure of substrate-bound human mtRNase P, which reveals the architecture of the *hs*PRORP active site in a pre-catalytic state with the RNA positioned for cleavage (Figure 5D) [59]. Its arrangement is reminiscent of the active site configurations of
human exonuclease 1 (hsExo1) (Figure 5E) and the Klenow fragment, suggesting that protein-only RNase Ps employ a similar two-ion mechanism as these exonucleases, as previously proposed [39, 83–87]. The metal ion at site 2 (M2), together with a proximal aspartate (D499 in hsPRORP), may position an activated water molecule opposite of the scissile P–O bond for nucleophilic attack. The metal ion at site 1 (M1) may stabilise the transition state by coordinating the oxygen atom of the scissile P–O bond. The recent crystal structure pbHARP reveals a similar configuration of catalytic residues and coordination of two metal ions in the active site, suggesting that HARPs likely share the two-ion catalytic mechanism [65]. Notably, RNP RNase Ps have also been proposed to use a two-metal catalytic mechanism, which conceptually resembles that described above [29, 35, 88–91]. Taken together, the structures of PRORPs and HARPs indicate...
that protein-only RNase P enzymes use an ancient catalytic mechanism employed by many functionally disparate nucleases.

Concluding remarks
PRORP enzymes represent a striking example of convergent evolution, as they carry out an identical reaction as their ribzyme-based counterparts but are evolutionarily unrelated. Since their discovery several decades ago, a combination of structural and functional studies has resulted in a basic molecular understanding of protein-only RNase P enzymes. These studies have revealed a surprising diversity of molecular architectures among these enzymes, as well as a conceptually unique substrate recognition and binding mode used by msPRORPs. Despite these differences, however, their catalytic mechanism appears to be highly conserved. Nevertheless, a high-resolution structure with substrate bound in the active site is thus far only available for the eukaryotic mitochondrial msPRORP [59]. As a consequence, many mechanistic assumptions for ssPRORPs and HARPs are based on extensive biochemical data as well as structural and evolutionary comparison. Therefore, the next important step will be to obtain molecular snapshots of ssPRORPs and HARPs in complex with their respective substrates in the active site in order to verify these hypotheses. Furthermore, the biological functions of individual PRORP enzymes will need to be further investigated. The now available structural and mechanistic data will provide a framework for future studies aimed at answering the many open questions (see Outstanding questions) about protein-only RNase Ps.

Declaration of interests
The authors declare no conflicts of interest.

References

Outstanding questions
What is the atomic structure of full-length substrate-bound ssPRORPs?
What is the architecture of a dodecameric, substrate-bound HARPs complex?
What are the in vivo substrates of protein-only RNase Ps, especially in organisms that contain both RNP- and protein-only RNase Ps?
Do additional, yet undiscovered types of protein-only RNase P enzymes exist?
How are protein-only RNase Ps regulated?
84. Orans, J. et al. (2011) Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclelease family. Cell 145, 212–222