CHAPTER 5

An Introduction to RNase P

James A. Pannucci¹, Elizabeth S. Haas, and James W. Brown
Department of Microbiology, North Carolina State University,
Raleigh, NC, USA; ¹Present Address: Life Science Division, Los
Alamos National Laboratory, Los Alamos, New Mexico, USA

Background: Ribonuclease P is the enzyme responsible for the specific removal of 5' leader sequences from precursor transfer RNAs in all cells. RNase P enzymes are generally ribonucleoproteins, and in Bacteria and some Archaea, the RNA subunits alone are catalytically active in vitro, i.e., they are ribozymes. RNase P is an excellent system for investigating RNA structure and function, and the evolutionary processes that create, and continue to recreate, functional RNAs.

Purpose: We want to introduce general readers to RNase P structure, function, and evolutionary variation. Areas of current research efforts are introduced, particularly those leading to three-dimensional models/structures of RNase P components and the potential use of RNase P in gene therapies.

Techniques: A set of basic experimental methods for cloning, assaying, manipulating, and analyzing RNase P RNAs are described.
1. HISTORY

Ribonuclease P is the enzyme responsible for the specific removal of 5’ leader sequences from precursor transfer RNAs in all cells and organelles that synthesize tRNA (2,3,4,6,63,64). RNase P is composed of a highly-structured RNA (250-450 nucleotides) and by protein subunit(s) that range from a single 14-kDa polypeptide in Bacteria, to a large, multi-protein complex in the yeast nucleus (13,17,45,63). Previous work with bacterial RNase P has demonstrated that the RNA is the catalytic subunit. Physiological conditions require the presence of both subunits for efficient pre-tRNA processing, but under appropriate in vitro conditions (increased ionic strength), the RNase P RNAs from Bacteria and some Archaea are catalytically active without the RNase P protein (29,65).

“RNA-alone” activity in vitro has not been demonstrated by RNase P RNAs from eukaryotic nuclei (67), organelles (26,28,53,54,70), and several Archaea (50,65), but the RNase P RNAs are still presumed to be the catalytic subunit based on a shared and highly conserved set of functional core RNA components (18). The tertiary structure of the RNase P RNA, formed by the interaction of conserved and variably present secondary structural elements, is undoubtedly the key to catalytic activity.

In vitro, the RNase P RNA is bound by its protein subunit(s) that are responsible for shielding the electrostatic repulsion of the RNA backbone, thus enabling RNA tertiary folding (68) as well as facilitating the recognition of the substrate (49). RNA-alone catalysis requires structural elements to be arranged in a proper tertiary conformation so that the global architecture is sufficient to correctly position both the enzyme and substrate, but even if all key structural features are present, incompatible global arrangements can prevent enzymatic function. The in vitro phenomenon of RNA-alone catalytic activity, and the conditions required to produce that activity, demonstrate the ability of an RNase P RNA to self-fold and also demonstrate that it contains all the necessary structural features to bind pre-tRNAs and cleave 5’ leaders (29).

In this review, the structure and function of the RNase P RNA will be considered in light of the common aspects of RNase P, its role in catalysis, and how it is affected by various present structural elements and the RNase P protein subunit. Readers may wish to examine specific secondary structures in addition to those presented here by viewing the RNase P Database, on the World Wide Web at http://www.mbio. ncsu.edu/RNaseP (8).

1.1 Substrate recognition and cleavage site selection

The well-understood secondary and tertiary structure of tRNA has been applied to understanding the substrate recognition and cleavage site selection requirements of RNase P. In vivo, RNase P processes all the pre-tRNAs of a given cell or organelle. In vitro, RNase P will process pre-tRNAs from any species (2,21,64). The RNase P interaction with substrate depends on the highly conserved features of the mature tRNA domain rather than aspects of the precursor that may vary with each specific form (16). Substrate interaction experiments have identified the acceptor stem and the TΨC stem-loop as the critical recognition structures (38,39) (Figure 1). The dihydouracil (DHU) and anticodon stem-loops serve as additional, but not essential, anchors (16,58).

The acceptor stem of almost all pre-tRNAs contains a seven-base-pair helix that appears critical for correct cleavage site selection (42,75,76). RNase P from Xenopus laevis apparently measures the stem length to find the appropriate cleavage site (15). The insertion of two base pairs in the acceptor stem of a yeast pre-tRNA leads substrate alters the cleavage site by shifting it two base pairs 5’ to maintain the seven-base-pair acceptor stem. These changes also affect the catalytic efficiency of substrate processing (alters $K_{cat}$ and $K_m$). Analysis of RNase P cleavage of circularly-permutated substrates (circularly-permutated pre-tRNAs have 5’ and 3’ ends generated in unusual locations after circularization and hydrolysis of substrate RNA) shows that interruptions in the pre-tRNA sugar-phosphate backbone are not generally tolerated in the RNA acceptor stem or in the TΨC stem-loop (16). In addition, the Escherichia coli RNase P (both the RNA-alone and the holoenzyme) has been shown to require specific nucleotides to be present near the cleavage site of the pre-tRNA (75).

3’-NCCA tails are present in all tRNAs. E. coli and many other species have the CCA sequence encoded in most tRNA genes, but have the CCA added as needed by a separate enzymatic system. A CCA tail is not essential for processing by RNase P but it does greatly influence rate and cleavage site positioning (40,43,48,62). E. coli RNase P RNAs that have mutations in the P15/16 loop region that base pairs to the tRNA CCA tails exhibit aberrant cleavage (43). The P15/16 loop region of RNase P RNAs is not universally conserved. In the Archaea, many of the RNase P RNAs without the P15/16 loop GGU sequence are not catalytically active in vitro (65). The idea that the protein subunit(s) of these RNase P enzymes compensate for the absence of CCA addition to the RNase P RNA is appealing. However, there are examples of bacterial (e.g., cyanobacteria and Clostridium) as well as archaearial RNase P RNAs (e.g., Thermococcus cell) that do not have P15/16 loop sequences complementary to CCA, but perform in vitro catalysis accurately without a protein subunit (35,65).

![Figure 1: Pre-tRNA* from R. meliloti This commonly-used RNase P substrate is synthesized by T7 transcription in vivo from plasmid pDW118 cloned with AesI (70). The substrate has the mature 3’ terminus but no base modifications. The 5’ leader is not the native sequence; it is derived from cloning vector sequence.](image-url)
Pre-tRNAs can be covalently cross-linked to various positions on the RNase P RNA. The cross-links are dependent on the tertiary arrangement of both the enzyme and substrate RNAs, but structural conformation of the RNase P RNA is not the only important consideration. There are specific Watson-Crick base pairs formed between the enzyme and substrate RNAs. The CG sequence on the 3' side of the P15V16 loop of the E. coli RNase P RNA (Figure 2) base pairs with the CC sequence in the 3' CCA-tail of the substrate. Also possible is the extension of the base pairing to include the following base in RNase P RNA (usually a U) if there is a purine in the rRNA 'discriminator' position (the generally-unpaired base preceding the 3' CCA). The presence of the base pair interactions have been corroborated by compensatory mutations (43). The end of the acceptor stem of rRNA is thought to be a Mg** binding site and important for RNase P cleavage activity (73). The formation of base pairs near the RNase P cleavage site of the substrate may position and expose the cleavage site by unwinding part of the acceptor stem (43).

Full-length pre-tRNAs are not the only substrate recognized by RNase P RNAs and holoenzymes. The 4.5 S RNA and the Signal Recognition Particle (SRP) RNAs of Archaea and Eukaryotes are natural, non-pre-tRNA RNase P substrates that mimic the structure of the tRNA acceptor stem that is stacked on the WC stem-loop (Figure 1) (44-47). The 10 S RNA and certain plant viral RNAs have tRNA structural-mimics near the ends of RNA sequences that are processed by RNase P (31,47,55). Two hypotheses of RNase P processing in plant viral RNAs have been suggested: one is that RNase P is associated with a host defense function, and the other suggests that RNase P activity is part of a multi-step RNA editing event. The 10 S RNA tRNA-like domain is likely processed by RNase P to facilitate its interaction with the ribosome and its role in abortive polypeptide decomposition.

A natural substrate for RNase P has been identified in Drosophila (41) after infection by a 5 kb retrovirus-like RNA. In all other instances of retrovirus infection, full-length, mature tRNAs are used as primers for minus-strand reverse transcription, but in the Drosophila retrovirus-like particle infection, a fragment of the mature initiator tRNA that is corresponding to the 5'-most 39 bases is used to prime reverse transcription. Although convincing evidence is still forthcoming, tRNA-Met is thought to assume an alternate conformation that is an unusual RNase P substrate. The alternate fold most likely arranges an acceptor-arm-like helix and DHU stem loop in a coaxial stack recognized by RNase P; cleavage occurs between bases 39 and 40 of the mature tRNA-Met after the initial RNase P maturation of the 5' end occurs. The 5'-most 39 bases, freed from the 3' section of the molecule, bind the Drosophila retrovirus-like RNA and prime reverse transcription.

1.2 Catalytic mechanism

RNase P catalyzes a direct hydrolysis rather than a transphosphorylation reaction that occurs in some other ribozymes (73). Beebe and Fierke (6) proposed a step model of RNase P RNA catalysis based on the RNase P RNA and pre-tRNA from Bacillus subtilis. The cycle begins with rapid binding of pre-tRNA by the RNase P RNA followed by cleavage of the correct phosphodiester bond.

Dissociation of products occurs in a kinetically-favored pathway, the 5' fragment leaves first followed by the rate-limiting tRNA dissociation. The final step is the formation of a second conformation of RNase P RNA that is not stable and binds pre-tRNA more slowly.

A photo-cross-linking assay has been used to determine the affect of metal ions on RNase P substrate affinity (72). Combined with data from RNase P activity assays, cross-linking of bacterial RNase P RNA to various pre-tRNAs suggests that monovalent ions primarily reduce electrostatic repulsion and facilitate substrate binding. The ionic radius of different monovalent ions had an affect on the RNase P RNA. Cs+, K+, NH4+, and Rb+ have ionic radii of 1.3-1.7 angstroms. Ions with radii outside that range do not have a simple correlation between binding and cleavage of substrate. Specifically, Na+ has an ionic radius of 0.95 angstroms and is threefold less effective than NH4+ in binding substrate at a concentration of 1 M, but at 1 M Na+ affords a threefold faster catalytic rate (72).

Various ions were also tested for their affect on the in vitro archaeal RNase P RNA reaction (63). The monovalent ions, K+, Na+, and NH4+ stabilized the Methanococcus thermoautotrophicum strain A4 RNase P RNA (Figure 3), and cyt-P** could enable pre-tRNA processing similarly to Mg**. As with the bacterial RNase P RNA, cyt-P** was more effective in the archaeal RNase P RNA reaction at lower concentrations than Mg**, but it also increased RNA degradation. Polynucleotides, such as spermine and spermidine, cannot compensate for the high ionic strength requirements of either the bacterial or archaeal RNase P RNA-alone reactions (65,72). In general, the extreme salt requirements and high Km of catalytically active archaeal RNase P RNAs resemble effective bacterial RNase P RNA mutants, but archaeal RNase P RNAs are clearly homologous to bacterial
RNase P RNAs share the core structural components, and respond in a similar manner to various ions. Divalent cations have three apparent functions that are related directly to catalysis: facilitation of the deprotonization of water molecules, polarization of a phosphoril oxygen to increase electropositivity of the corresponding phosphorus, and coordination of the phosphoril oxygen to stabilize the transition state. Mg** ions are required for cleavage, but not for the formation of enzyme-substrate complexes. Reaction velocity increases with Mg** concentration until affinity between product tRNA and the RNase P RNA slows product release and lowers reaction rates in multiple-turnover reactions. At least three Mg** ions are required for optimal catalytic activity. The 2' hydroxyl in the phosphodiester linkage cleavage site appears to bind one of the Mg** ions (73). Ca** allows RNase P cleavage of substrate to occur, but at slower rates than Mg**. Mn** is also able to catalyze RNase P reactions. Mn** has a greater nucleotide affinity than Mg** and is more effective at lower concentrations, but the denaturing effect causes tight binding of Mn** to nucleotide bases and phosphates, thus reducing substrate binding and catalytic rate (78,72).

1.3 Variable elements, conserved core, and consensus structures

Although in vitro catalytic activity has only been demonstrated by bacterial and certain archaeal RNase P RNAs (65), the RNA subunit is presumed to be the catalytic entity in all cases where it is present (a protein RNase P has been proposed in certain cases where an RNA subunit cannot be found (26,70)). Structural conservation, in particular of the catalytic core of the molecule, implies functional conservation as well (18).

For an RNase P RNA to be active in vitro, all components have to be arranged in a catalytically competent global configuration. In vitro RNA-alone reactions of natural or synthetic RNase P RNA variants provide information about the function of various structural components. The presence or absence of non-core structural elements (i.e., structures that are naturally present in some RNase P RNAs but not others) can alter the biochemical properties of the RNase P reaction in vitro in ways that can be interpreted. For example, an RNase P RNA that is inherently unstable and fails to be properly stabilized by salt will have a low affinity for substrate with the possibility of poor or no in vitro catalysis (65).

Five conserved sequences (Conserved Regions 1-5; Figures 2 and 3) have been proposed as having specific homologies in all cellular RNase P RNAs (18). However, the secondary structures of these RNAs are much more highly conserved than are their sequences. Much of this structural identity has been identified by comparison analysis (52,78), but confirmation has been provided by physical experiments designed to probe RNase P RNAs for structural features (23,37,57). Type A RNase P RNAs, such as the RNase P RNA from E. coli, are the most common and represent the ancestral form. The type B RNase P RNA occurs in Gram positive bacteria such as B. subtilis (Figure 2) and are a derived structural form (34). Although no single RNase P RNA contains all of the observed helical variations found in the various instances of bacterial RNase P RNAs, each contains a common core of sequence and structural elements. All bacterial RNase P RNA structures contain the helices P1-P5, P7-11, and P15. Helices P12 and P18 are present in >98% of the known bacterial sequences (32,71).

Variable helical elements often protrude from regions of the RNase P RNA core and are most highly conserved at their bases. Terminal loop sequences are also often conserved because they can serve as structural anchors via tertiary interactions with nucleotides elsewhere in the molecule (35). The helices P8, P9, P13, P14, P16-17, and P18 of bacterial RNase P RNAs are examples of helices with conserved bases and terminal loops. P9 originates in the cruciform (P7-P10) and its terminal loop, L9, docks with nucleotides in P1 (37). The loop of P8 docks with P4 (57). P13 is known to interact with P12, although what this interaction might be is not known (12). P16-17 is located distal to the P15 loop region, which is part of the minimal bacterial consensus structure, and its terminal 'loop' interacts with sequence between P5 and P7 to form P6 (50). In the type B RNase P RNAs, P6, P16, and P17 are replaced by P5.1 and P15.1, which are also not part of the minimal consensus. P18 and P14 extend from highly conserved regions in RNase P RNA, and their terminal loops interact with adjacent base pairs in P8 and are important stabilizing structures (12).

The nucleotide sequences of RNase P RNAs range from 100% to less than 25% identical. Variation often occurs as the insertion or deletion of non-essential structural elements and in the sequences of these helices (35). Non-essential structural features are not absolutely required for catalysis, but are often important as elements of structural stability.

Under optimal ionic conditions, the native Chlorobium sp. RNase P RNAs have a Kd similar to the native E. coli RNase P RNA. The Chlorobium RNase P RNAs do not contain P18, a structure that is highly conserved in helical length and loop sequence among other Bacteria. When P18 is deleted from the E. coli RNase P RNA (E. coli ΔP18), the requirement for monovalent salt is increased (32). Mutant bacterial RNase P RNAs and the catalytically active archaeal RNase P RNAs (65) require increased ionic conditions for structural stabilization.
The absence of P18 from the Chlorobium RNase P RNA indicates that it is not essential for catalysis. How the absence of P18 is compensated for in Chlorobium is not known, but bacterial RNase P RNAs lacking P18 or with an unusual P18 structure also contain elongated P8 helices, which is consistent with the known interaction between these helices.

The presence of helix P12 is highly conserved in RNase P RNAs, but is absent from the RNase P RNA of the bacterial species, M. fermentionis. The absence of P12 from the catalytically active M. fermentionis RNase P RNA signifies that it is not an absolute requirement for catalysis (71). The Km of M. fermentionis RNase P RNA is higher than the native E. coli RNase P RNA at 1 M NH₄⁺, but the Km was similar at 3 M NH₄⁺. The impact of P12 as an important structural component was observed in an E. coli RNase P RNA P12 deletion mutant (E. coli P12), where its absence caused a tenfold decrease in catalytic activity at 1 M NH₄⁺, but only a twofold decrease at 3 M NH₄⁺ (71).

The RNase P RNAs Min 1 (79) and Micro P (71) have variably-present regions deleted to produce 'minimal' RNase P RNAs. Min 1 is a simplified hybrid of the E. coli and B. subtilis RNase P RNAs, while Micro P was made by deleting helices from the structure of the M. fermentionis RNase P RNA that are not found in type A RNAs. Both Min 1 and Micro P are catalytically active in vitro, but are structurally compromised and require very high ionic strength for activity. The minimal RNase P RNAs show that core catalytic elements will process pre-tRNAs if satisfactorily positioned, and that the variably present structural elements of native RNase P RNA stabilize the global arrangement of the molecule.

Archaeal RNase P RNAs are similar in structure and sequence to the bacterial RNase P RNAs (33). Some of the archaeal RNase P RNAs do not have a loop between P15 and P16, a highly conserved region in bacteria that interacts with the CCA-tail of pre-tRNAs. Interestingly, many of the archaeal RNase P RNAs that have catalytic activity in vitro have the P15/P16 loop with the GUU sequence that is present in most bacteria (65). The archaeal secondary structure is likely to be refined in the future because there are considerably fewer potential sequences than seen in bacteria (56 versus 182) (8).

The collection of eukaryotic RNase P RNA sequences is steadily increasing in size, but a comprehensive secondary structure model has not been reported. Based on a vertebrate nuclear RNase P RNA structure constructed from an alignment of 17 sequences, the eukaryotic RNase P RNAs apparently lack P15-P18, but they do have a large internal loop in P3 (Figure 4) (67). The P15-P18 region is also absent in a recently reported secondary structure of the Schizosaccharomyces pombe RNase P RNA (18). The loss of P15-P18 and its associated loop changes the RNA secondary structure with the benefit of meaningful comparative analysis. Organellar RNase P RNAs are equally difficult in physical experiments; their unique sequence and structural characteristic make identification a challenge that becomes no less difficult when confronted with the possibility of an all-protein RNase P as postulated for the human mitochondria (70) and vascular plant chloroplasts (26).

The plastid (cyanide) from the protist Cyanophora paradoxa contains an RNase P RNA (Figure 5) that appears to be inactive in vitro without its protein component. This RNA is predicted to have a secondary structure like that of bacteria and can be made functional when reconstituted with the RNase P protein from Synechocystis (a cyanobacterium) but not with the RNase P protein from E. coli (66). This is the first reconstruction of RNase P activity from organellar and bacterial subunits and supports the idea that the cyanide RNase P RNA has a similar structure and function to the cyanobacterial RNase P RNA from which they arose.

1.4 The RNase P holocomplex

Bacterial RNase P contains more RNA (120 kDa) than protein (14 kDa), which is reflected in the buoyant density (1.55 g/mL) of the E. coli RNase P (52) in Cs₂SO₄. Eukaryotic RNase P has lower

---

Figure 4. The S. cerevisiae and human nuclear RNase P RNA secondary structures (67,77). Helices with homologous positions in the bacterial RNAs are labeled P1-P19 according to Reference 32.
buoyant densities (1.34 g/mL for the wheat RNase P) that correspond to a much larger protein:RNA ratio (4).

Instead of a single polypeptide protein subunit, as in bacteria, the yeast nuclear RNase P protein is comprised of nine protein subunits (17). Although archael RNase P's have been purified (22,59), the number and identity of protein subunit polypeptides has not been determined. Buoyant density measurements indicate mixed possibilities. For crenarchael organisms, like Sulfolobus acidocaldarius (density = 1.27 g/mL) there is a high protein:RNA ratio, thus suggesting large or multiple polypeptides. Euryarchael organisms, such as Haloferax volcanii (density = 1.61 g/mL) (52), have a higher buoyant density that suggests a lower protein:RNA ratio and the likelihood of few or small polypeptides.

The single polypeptide bacterial RNase P protein appears to be unrelated in sequence to any of the nine polypeptides associated with the yeast nuclear RNase P (17,46). The yeast proteins were isolated from Saccharomyces cerevisiae and several are shared by the mitochondrial ribosomal processing ribonuclease (RNase MRP). In the case of the archaea, three complete genome sequences are now available, but the protein component(s) of the RNase P have yet to be identified.

Yeast mitochondrial RNase P activity is present and micrococcal nuclease sensitive. Although the identity of the RNA has not been determined, the buoyant density of the enzyme is 1.28 g/mL, suggesting a high protein:RNA ratio (28). The yeast mitochondrial RNase P protein is the product of the RNMI2 gene; this protein apparently also has an essential nuclear function not related to RNase P (20,56). Human mitochondrial RNase P has a density of 1.23 g/mL, indicating an enzyme with a high protein content. This result, combined with the protein's resistance to rigorous treatment with nucleases, may suggest that the RNase P is not a ribonucleoprotein but rather a protein enzyme (76).

The biochemical analysis of the function of the RNase P protein was an obvious curiosity after demonstration of the RNase as the catalytic subunit. Early experiments compared the B. subtilis RNase P RNA reaction (100 mM NH₄Cl, 100 mM MgCl₂) with the holoenzyme reaction (100 mM NH₄Cl, 15 mM MgCl₂) and demonstrated that the RNA and the holoenzyme had a similar Kₘ. The addition of RNase P protein increased kₚ by 20-fold, which led to the conclusion that the protein facilitated product release in addition to shielding electrostatic repulsion (68). Later experiments under different experimental conditions (100 mM NH₄Cl, 10 mM MgCl₂ in the RNase alone reactions) and using the B. subtilis RNase P RNA and holoenzyme have provided an alternate conclusion. The B. subtilis RNA P holoenzyme appears to have greater affinity for the precursor tRNAAP than it does for the mature molecule, thus, the major role of the protein is to facilitate the discrimination between the

---

**Figure 5.** The *C. paradoxa* and *P. purpurea* plastid RNase P RNA secondary structures (5,69). Helices with homology to the bacterial RNAs and have not been determined experimentally.

**Figure 6.** The *R. americana* and *A. nidulans* mitochondrial RNase P RNA secondary structures (51,56). Helices with homology to the bacterial RNAs are labeled P1-P19 according to Reference 32.
tRNA precursor and product, rather than the stabilization of the tertiary structure of the RNA (49).

2. RESULTS

2.1 RNA-alone catalysis by bacterial and archaeal RNase P RNAs

The hallmark discovery in the study of RNase P was that bacterial RNase P RNAs possess the necessary structural components to be catalytically active in vitro without protein. The RNase P RNA, by itself, can perform each step in the RNase P reaction cycle. RNA-alone activity has been observed for all tested bacterial RNase P RNAs (11,29,32) and also for archaeal RNase P RNAs from some of the Euryarchaeota: the methanobacteria, halobacteria, and thermoacidophiles (65) (Figure 7).

The natural forms of other catalytic RNAs perform only a single intramolecular reaction, RNase P is unique because it is a trans-acting catalytic RNA, an enzyme that reacts with multiple substrates. The demonstration of catalytic activity by an RNase P RNA in the absence of protein is a significant achievement because it shows that RNA can be truly enzymatic and it provides a simple model to examine the relationship between structure and function in RNA (63).

2.2 Sequence/structure dataset

Secondary structural refinement and tertiary model construction has been made possible by the continual collection of RNase P RNA sequences. Sources of RNase P RNA genes are pure cultures, sequences of complete genomes, and isolation from environmental samples. Obtaining additional sequences is critical because the discovery of unique variations may represent evolutionary intermediates or reflect an evolutionary change not observed in RNase P RNAs from other species. In addition, the statistical methods of comparative analysis decrease in uncertainty and increase in resolution as the sample size increases.

2.3 Three-dimensional models

Structural arrangement is the critical determinant of catalytic activity. Models of the three-dimensional structure of RNase P RNA are devised to gain insight into global structure and to propose hypotheses about how the molecule functions (Figure 8) (37,57,61). Three-dimensional models have been built and refined based on the combination of comparative analyses and cross-linking data. Accurate three-dimensional models represent an important achievement in the study of RNase P RNA because they allow visualization of the catalytic core of the molecule and approximate the location of the structure-stabilizing peripheral elements. Construction of a tertiary model requires the identification of core versus peripheral components, but also, and more significantly, where the various structural features are located in space and in reference to one another.

2.4 Individual structural components

Tertiary models of the RNase P RNA structure predict the arrangement of structural features but not necessarily their function. A notable achievement in the study of RNase P RNA has been the dissection of structure and function.

The stabilizing role of bacterial RNase P RNA helices P18 and P12 have been discussed previously. Equally impressive work has been performed to elucidate Mg++ binding sites and the substrate tRNA, CCA-tail recognition functions of the P15/16 loop of type A and P15 loop of type B, bacterial RNase P RNAs (27, 43, 62). Understanding the function of individual structural features is in turn valuable as a constraint in tertiary model construction. Computer generated models of the bacterial RNase P RNA P15/16 loop region (24), and the Mg++ binding P10/11-P12 region of yeast RNase P RNA (83) had to account for function when predicting their global configuration.

2.5 An RNase P protein crystal structure

The recent report of the B. subtilis

![Figure 7. Phylogenetic tree based on 16S-like ribosomal RNAs (7).](image_url)

![Figure 8. Model of the tertiary structure of E. coli RNase P RNA (57). Balls represent nucleotides and are colored on the basis of sequence variation; highly conserved nucleotides are dark, highly variable nucleotides are light. This view shows the pre-RNA binding surface. Helices that are clearly visible in this angle are labeled P1-P17 according to Reference 32.](image_url)
RNase P protein crystal structure (74) added another interesting feature to the already remarkable enzyme. The protein subunit of RNase P has an unusual left-handed βαβ crossover connection rather than the right-handed crossover connection typically seen between β strands. The unusual RNase P protein topology is shared with the ribosomal protein S5 and the ribosomal translocase elongation factor G and may represent evolution from a common RNA binding ancestor. The RNase P protein also has a large central cleft and a metal-binding loop, which, together with the left-handed βαβ, constitute three possible RNA binding sites. The discovery of a rare RNA binding motif shared by certain ribosome-associated proteins is a significant achievement that will contribute to the unraveling of the evolutionary history of RNase P.

2.6 RNase P-based gene therapy

The preliminary work with RNase P as an agent of gene therapy is a significant achievement (1,25,82). The most promising approach to RNase P based gene therapy is the design of external guide sequences (EGSs). An EGS is an oligoribonucleotide that is in part complementary to the sequence of a target RNA (Figure 9) (25). The annealing of the EGS to a target RNA (e.g., mRNA or viral RNA) results in the formation of a pre-RNA-like structure (80,81). EGSs require bases that are complementary to the target RNA and a 3' CCA for optimal cleavage activity. Target RNAs become the 5' portion of an RNase P substrate, and cleavage is directed at a particular base in the target RNA. Extra sequence of target RNAs are tolerated so long as the structural requirements of the RNase P substrate 5' and 3' ends are met. Endogenously RNase P will recognize the trans-assembly substrate and cleave the target RNA, rendering it biologically inactive. By using EGSs, any RNA is a potential target (25).

3. CONCLUSIONS

The study of structure and function encompasses nearly all aspects of RNase P. Catalytic requirements, kinetic parameters, secondary structures, and tertiary models are all dependent on the conservation of a structural core that is created by a variety of sequences and secondary structures and is influenced and stabilized by variably present peripheral elements. The RNase P protein subunit is required in vivo but not necessary for bacterial and some archael RNase P RNAs in vitro. In vitro RNA-alone activity indicates that the RNase P RNA contains the essential elements for substrate recognition and catalysis, and they possess the ability to self-assemble into a catalytically competent tertiary structure. In instances when RNase P RNAs are not catalytically active in vitro, a more complete understanding of the RNase P protein combined with increasingly detailed tertiary models should clarify the discrepancy between lack of activity and the function of a conserved structural core.

4. FUTURE TRENDS

Future trends in RNase P research are likely to involve the increased study of the RNase P protein subunit, both alone and as part of the RNase P holoenzyme. The recently determined structure of the B. subtilis RNase P protein (74) is the first look at the structure of an RNase P protein and how that structure relates to its function. With continuous improvements in crystallization methods for RNA, it seems inevitable that the structure of the RNase P RNA (or holoenzyme) will eventually be determined by X-ray diffraction. Cryo-electron microscopy has already been applied to the structural study of ribosomal subunits with great success. Combined with increasingly sophisticated imaging methods, improved experimental techniques will produce high-resolution, detailed images of RNase P perhaps taking advantage of naturally-occurring structural variations between species to identify structural elements in the three-dimensional reconstructions.

The details of structure will contribute to the understanding of function. Armed with this knowledge, RNase P applications such as gene therapy can be advanced to the point of predicting or designing function based on knowledge of structure. The progression of RNase P based gene therapy is a good example of the initial efforts in this direction. Understanding the structure of the RNA, in particular the RNA:RNA enzymesubstrate interaction, led to the design of effective EGSs.

Future experiments may resolve issues such as the instances of RNase P where an RNA has not been identified. If it is true that RNase P activity can be accomplished by an all-protein system, it would be interesting to compare the novel catalytic mechanisms to those understood for the catalytic RNase P RNA. This is an opportunity for those interested in early biochemical evolution; the so-called “RNA World” hypothesis demands the take-over of RNA structure and their functions by polypeptides, but clear examples of such takeover in extant life, e.g., the occurrence of a specific functional (or structural) role carried out by an RNA in one organism and carried out by a protein in another, are not known.

5. TECHNIQUES

5.1 Cloning RNase P genes

The difficulty in cloning RNase P RNA genes has been greatly reduced by the accumulation of sequences from different organisms. Hybridization of labeled probes to bacterial colonies with plasmids containing fragments of genomic DNA has been used extensively to obtain RNase P RNA genes (10,11,32,
34). The collection of sequences, combined with comparative analysis, identified highly conserved regions of RNase P RNAs that were specific for particular evolutionary groups (10,33-35,67). PCR amplification with domain-specific primers often yields only partial RNase P RNA genes. The primer annealing sites roughly correspond to the RNase P RNA helix P4, and omit P1, the 5' portion of P2, and P3 from amplification. Methods for obtaining the sequences upstream and downstream of the domain-specific primers are available when working with pure, cultivated species, but are not feasible when working with genes obtained from environmental samples or mixed cultures. PCR primers can also be designed based on the sequence conservation in P1 at the 5' and 3' ends of RNase P RNAs (32,33). This narrows the evolutionary scope of gene searches, allowing the amplification of RNase P RNA sequences from additional species of a genus from which a representative full-length sequence is available.

DNA for use as template in PCR amplification experiments can be obtained from a variety of sources. Amplification of DNA from pure cultures is common, as is obtaining RNase P RNA genes by traditional methods, typically Southern blotting and colony hybridizations using genomic DNA preparations. Isolating DNA from prokaryotic species requires relatively simple extraction and precipitation procedures (10), and kits are now available that are generally quite useful. Complete genome sequences are a valuable source of RNase P RNA gene sequences, and the information is useful in obtaining sequences from related species.

PCR amplification from natural microbial populations is a technique that eliminates the need to isolate or grow individual species (12). An environmental sample, such as wastewater sludge or pond water, is collected and DNA is extracted. Specific PCR primers are then used to amplify a collection of RNase P RNA genes, which are then cloned. Amplification of novel genes from natural populations of organisms often produces sequences from unidentified species, but evolutionary relatedness can be determined to various degrees based on phylogenetic analysis of the RNase P RNA sequences. This approach is advantageous because it generates many more sequences than could be obtained from growing pure cultures, especially since the majority of organisms in any particular sample cannot be readily cultured.

5.2 RNase P activity assays

The RNase P RNA activity assay is used to test function (Figure 10). Monovalent and divalent cation concentrations are critical factors when designing an RNase P RNA activity assay. NH₄⁺ and Mg²⁺ are preferred cations, but activity has also been observed using other ions (65,72). Activity assays can be attempted using an RNase P RNA from any source, extracted from cells or synthesized in vitro from cloned genes, but in vitro RNA-alone cleavage of pre-rRNA has only been observed in bacterial and some archaeal RNase P RNAs (35,65).

Bacterial RNase P RNAs are catalytically active under a wide range of ionic conditions, but 1 M NaCl and 10 mM MgCl₂ produces optimal catalysis in vitro for E. coli and B. subtilis RNase P RNAs. Mutant bacterial RNase P RNAs often require increased ionic concentrations, presumably to stabilize poorly formed tertiary structures (63). High salt concentrations have also been used to test the catalytic activity of bacterial RNase P RNAs at elevated temperatures (up to 75°C), where increased stability is useful to prevent thermal disruption; the E. coli RNase P RNA is far more active in reactions containing 3 M NaCl at 65°C than in standard conditions (11). The minimal ionic conditions needed to observe activity in the catalytic archaean RNase P RNAs is approximately 1-1.5 M ammonium acetate and 100 mM MgCl₂, but optimal conditions are much higher: 3-4 M ammonium acetate and 300-400 mM MgCl₂ (65).

For the bacterial and archaean catalytically-active RNase P RNAs, optimal pH is approximately 8.0 and is achieved by including 30 mM Tris-Cl pH 8.0 in reactions. Reactions typically contain 32P-labeled pre-rRNA as substrate, synthesized by in vitro transcription from cloned tRNA genes, and are resolved using 8% (wt/vol) denaturing polyacrylamide gel-electrophoresis.

RNase P activity assays are also performed using the RNase P holoenzyme. In such cases, salt is held closer to physiological conditions (100 mM NaCl, 5-10 mM MgCl₂) (49,68). Because the RNase P protein subunit stabilizes the RNase P RNA effectively, RNase P holoenzymes can be assayed at higher temperatures and in lower salt conditions than the RNA alone. Reconstitution of the RNase P holoenzyme from RNA and protein is measured by determining the boost in activity provided by the protein when assayed at low ionic strength.

5.3 Circularly-permuted RNAs

A circularly-permuted RNA is a molecule with artificial 5' and 3' ends that has its native 5' and 3' ends incorporated into continuous sequence. Circularly-permuted RNase P RNAs (or substrates) are used in cross-linking and unimolecular enzyme-substrate complex experiments (19,37,61).

Figure 10. RNase P enzyme assay. RNase P or RNase P RNA (puriﬁed, notiﬁed RNase P from M. thermoautotrophicum in this example) are incubated in buffer with uniformly 32P-Labeled substrate pre-rRNA, and reaction products are visualized by autoradiography following electrophoresis in 8% polyacrylamide. 7.5 M urea gels. In this example, the experiment is a titration of ammonium acetate; optimal activity is seen at approximately 800 mM ammonium acetate. Control reactions contain either no added RNase P (negative) or RNase P RNA from E. coli.

controls + 0.1-1.0M ammonium acetate

trNA precursor

trNA product

leader fragment
An RNase P RNA or pre-tRNA with artificial ends allows the addition of a molecular tag such as the photoreactive compound azidophenacyl bromide to be placed at specific locations within the RNA sequence. A unimolecular complex has the substrate RNA incorporated into the RNase P RNA. Substrate RNAs are usually joined to the enzyme RNA in the appropriate substrate recognition region, e.g., the RNA can be joined by its 3' end to the loop between P15 and P16, where pre-tRNA 3' recognition occurs. This largely removes the binding step from the catalytic cycle and allows direct evaluation of catalysis.

5.4 Cross-linking

Physical methods such as X-ray crystallography and nuclear magnetic resonance (NMR) are used to generate tertiary structural models of protein molecules and small RNAs, but these techniques have not yet been successful with RNAs as large as RNase P RNA. Instead, ternary models of RNase P RNA are constructed using constraints from comparative data combined with intermolecular and intermolecular UV-light-induced specific cross-links (14,19,37,60). For example, cross-links demonstrated that the loop at the end of E. coli RNase P RNA helix P9 is adjacent to sequences in P1 (19). Because this is consistent with comparative data, any tertiary model must incorporate the proximity of these helices. Modeling studies that directly compared the tertiary arrangement of the E. coli and B. subtilis RNase P RNA revealed that the non-homologous structures P6/16/17 of E. coli RNase P RNA and P5/1-P15.1 interaction of B. subtilis RNase P RNA occupy the same space and probably serve an analogous function (19).

Experimental procedures are essentially the same whether an RNase P RNA is being cross-linked to itself (intramolecularly) or to a tRNA (intermolecularly). The RNA designated for the addition of the photoreagent is transcribed in the presence of guanosine monophosphorothioate (GMPS) that will incorporate into the nascent RNA only at the 5' position. RNA polymerase requires nucleotide triphosphates for chain elongation, therefore the GMPS will not be present at any other position. Circularly-permuted RNAs are employed to position GMPS on nucleotides located in the interior of RNAs.

The photoreagent commonly used in RNase P RNA cross-linking experiments is azidophenacyl bromide, which is specifically attached to the 5' shool group of GMPS. When RNA is exposed to UV light (302 nm), the azido moiety is converted to a nitrene that then inserts into nearby covalent bonds. Cross-linked products can be separated from non-cross-linked material on denaturing polyacrylamide gels. The cross-linked positions are determined by primer extension from oligonucleotides annealed to the target RNA; reverse transcriptase will elongate a DNA chain until it reaches the cross-link site (14).

Cross-linking of a nucleotide to an azidophenacyl photoreagent does not indicate the formation of base pairs or specific contact, but does demonstrate that the tagged nucleotide and the base it is cross-linked to are in proximity. The bond length of the photoreagent is 5-10 angstroms. Multiple cross-links are usually located near each other; however, when one tagged position has cross-links in distant regions of the target RNA it suggests close tertiary structural positioning, which is information vital to tertiary model construction.

5.5 Comparative analysis

The key to comparative analysis is conservation of higher-order structure despite changes in nucleotide sequence. Evolutionary changes of a nucleotide will correlate with compensatory changes elsewhere in the molecule that indicates interaction (usually direct) between these bases. The structure of an RNA can be determined by comparing homologous positions in a collection of RNA sequences for compensatory base changes. When coordinated changes in sequence occur in homologous RNAs from different species it is referred to as "covariation."

Covariation analyses begin with an RNA sequence alignment. Closely related sequences are easy to align because they will be similar in sequence and in sequence length. When aligning more distantly related RNAs, it is necessary to insert numerous gaps in the alignment that correspond to the existence of bases in one RNA that are not present in others. Once the RNA sequences are properly aligned, the construction of structural models is based on results from covariation analysis (39,35,36). Comparative analysis was used to construct secondary structures of many RNase P RNAs, and was recently used to refine the secondary structure and to identify tertiary interactions that lead to predictions of the tertiary structure of structural elements within bacterial RNase P RNAs (57). As mentioned above, photolytic comparative data is often combined with results of cross-linking experiments to produce a model that represents both comparative and physical constraints.

REFERENCES


