Ribozymes: the characteristics and properties of catalytic RNAs

N. Kyle Tanner *

Département de Biochimie Médicale, Centre Médical Universitaire, 1, rue Michel-Servet, 1211 Geneva 4, Switzerland

Abstract

Ribozymes, or catalytic RNAs, were discovered a little more than 15 years ago. They are found in the organelles of plants and lower eukaryotes, in amphibians, in prokaryotes, in bacteriophages, and in viroids and satellite viruses that infect plants. An example is also known of a ribozyme in hepatitis delta virus, a serious human pathogen. Additional ribozymes are bound to be found in the future, and it is tempting to regard the RNA component(s) of various ribonucleoprotein complexes as the catalytic engine, while the proteins serve as mere scaffolding – an unheard-of notion 15 years ago! In nature, ribozymes are involved in the processing of RNA precursors. However, all the characterized ribozymes have been converted, with some clever engineering, into RNA enzymes that can cleave or modify targeted RNAs (or even DNAs) without becoming altered themselves. While their success in vitro is unquestioned, ribozymes are increasingly used in vivo as valuable tools for studying and regulating gene expression. This review is intended as a brief introduction to the characteristics of the different identified ribozymes and their properties. © 1999 Federation of European Microbiological Societies. Published by Elsevier Science B.V. All rights reserved.

Keywords: Ribozyme; Catalytic RNA; Self-cleaving RNA; Self-splicing intron; Gene therapy; Review

Contents

* Tel.: +41 (22) 702 5508; Fax: +41 (22) 702 5502; E-mail: kyle.tanner@medecine.unige.ch

1. Introduction

Ribozymes, or catalytic RNAs, were first discovered in the laboratory of Tom Cech, at the University of Colorado, in 1982. His research team found that the ribosomal RNA precursor from *Tetrahymena thermophila* contained an intron, a nonencoding sequence that interrupts the gene, that was capable of excising itself, in vitro, without any protein or external energy source [1]. Shortly thereafter, Sid Altman's group, at Yale University, showed that the RNA component of RNase P, M1 RNA, from *Escherichia coli* was likewise able to process tRNA precursors without any protein factors [2]. Cech and Altman shared the Nobel Prize in Chemistry in 1989 for this work.

This seminal work ushered in a major research activity, and RNA catalysis was soon found to be widespread in nature, occurring in plants, bacteria, viruses, and lower eukaryotes. Although ribozymes are rare in vertebrates, one is found in humans. Seven distinct catalytic RNAs are identified in nature. This number is probably not exhaustive, and additional catalytic RNAs are bound to be found. It appears that the RNA components of some, perhaps many, ribonucleoprotein complexes may have catalytic activity; for example, highly deproteinized ribosomes can still catalyze a peptidyl transfer reaction [3,4], and nuclear mRNA splicing may be RNA-driven (see below). Moreover, new ribozymes are being generated, de novo, in the laboratory through combinatorial screening of randomized RNA sequences.

The fact that RNA can function as both a reservoir of genetic information and a catalyst of biological reactions has inevitably raised the supposition that RNA was the primordial biological molecule, and there is much written both for and against this idea – often based on the same experimental observations! The in vitro selection of catalytic RNAs, as well as the demonstration that ribozymes can catalyze polymerization reactions, has provided evidence for this ([5] and references therein). The difficulty of synthesizing discrete nucleotide building blocks under what is considered to be prebiotic conditions,

much less the difficulty of forming specific polymers from them, provides evidence against it (e.g., [6]). In the absence of a time machine, such speculations are bound to be rich sources of discussion for some time to come.

This article is intended as an overview of ribozymes, their structures, and their properties. However, a thorough discussion could easily fill a book; therefore, I will limit my presentation to more general properties of the molecules. Extensive reference will be made to recent reviews, and I will generally avoid citing articles that can be found through other, more recent, sources. As this review is intended to emphasize the structural features, as well as the biological functions, of ribozymes, I will not make extensive reference to the therapeutic or prophylactic aspects nor will I discuss strategies for delivering and expressing ribozymes in vivo. The reader is referred to some extensive reviews [7–9]. An excellent review of artificial ribozymes can also be found elsewhere [5].

2. General characteristics

Catalytic RNAs are broadly separated into two classes based on their size and reaction mechanisms (reviewed by [10–14]). The large catalytic RNAs consist of RNase P, and the group I and group II introns. These molecules range in size from a few hundred nucleotides to around 3000. They catalyze reactions that generate reaction intermediates and products with 3′ hydroxyls and 5′ phosphates (Fig. 1a). The small catalytic RNAs include the hammerhead, the hairpin (or paperclip), hepatitis delta and VS RNA. These molecules range in size from ∼35 to ∼155 nucleotides. They use the 2′ hydroxyl of the ribose sugar as a nucleophile, and they generate products with a 2′,3′-cyclic phosphate and a 5′ hydroxyl (Fig. 1b). The relationship between the size and the reaction mechanism of these molecules has raised intriguing questions about their origins and evolution. It may be that the reaction mechanism and the size of the large ribozymes are needed to

bring often very distal elements of the substrate into close proximity. The small, self-cleaving, RNAs are not faced with this constraint and perhaps this permitted them to evolve smaller catalytic centers. It remains possible, however, that the relationship between the size and reaction mechanism is simply fortuitous.

With one exception, all these RNAs catalyze reactions that modify themselves. Hence, they cannot be considered true enzymes or catalysts. The exception is RNase P, which processes the 5′ end of tRNA precursors. It is the only known example of a naturally occurring RNA-based enzyme. However, all these molecules can be converted, with some clever engineering, into true RNA enzymes that modify other RNAs *in trans* without becoming altered themselves.

Ribozymes increase reaction rates by up to 10^{11}-fold and have reaction efficiencies, k_{cat}/K_m, up to 10^8 M^{-1} min^{-1}, which is in the range for diffusion-controlled duplex formation between oligonucleotides [10]. While impressive, the rate enhancements provided by ribozymes are still $\sim 10^3$-fold less than those provided by protein enzymes catalyzing comparable reactions [15]. Moreover, ribozymes cannot compare with proteins as multiple-turnover enzymes, mostly because product release is so slow that the catalytic site of the ribozyme is easily saturated. This may be an inherent limitation of RNA enzymes, but it could also reflect evolutionary constraints, since ribozymes generally catalyze intramolecular, single-turnover, reactions in nature. An exhaustive comparison of the enzymatic mechanistics of protein and RNA enzymes has recently been made [15].

All known ribozymes have an absolute requirement for a divalent cation, which is generally Mg^{2+}. Some, notably within the large catalytic RNAs, require divalent cations for proper assembly of the tertiary structures as well. On this basis, catalytic RNAs are considered to be metalloenzymes, and a general two-metal-ion reaction mechanism has been proposed for the large catalytic RNAs, based on analogy with the properties of protein metalloenzymes [16]. The role of divalent cations for the small catalytic RNAs is less clear, but they are generally considered to be essential for catalysis (see [17] for alternative view).

3. General properties of introns

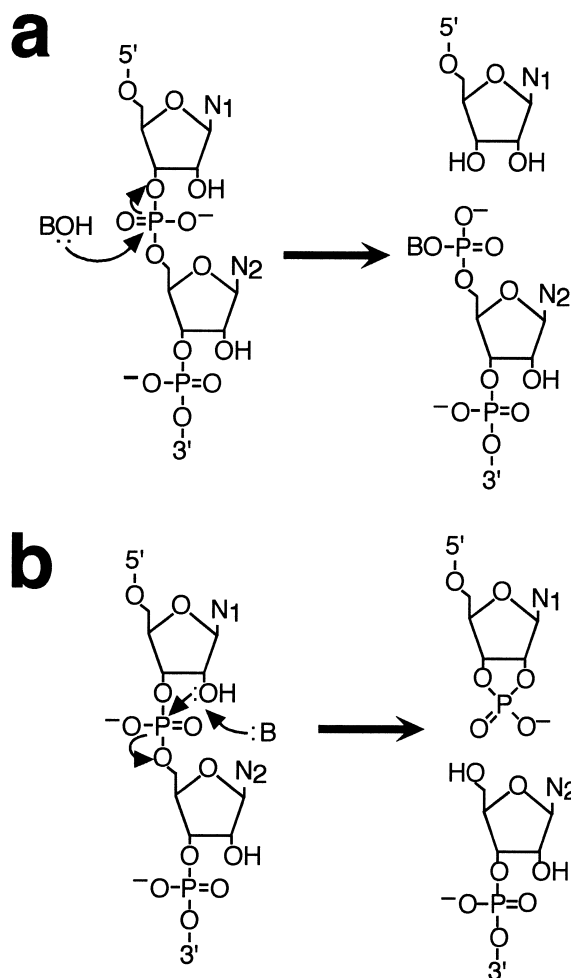Introns, or intervening sequences (IVS), are non-encoding sequences that interrupt the coding, exon,

Fig. 1. Reactions catalyzed by RNAs. a: The reaction mechanism of the large catalytic RNAs. The nucleophile BOH is the 3′ hydroxyl of a guanosine cofactor in the first step of group I splicing, the 2′ hydroxyl of a nucleotide, generally A, within the intron in the first step of group II splicing, and a water molecule in the RNase P-catalyzed reaction. b: The reaction mechanism of the small catalytic RNAs. The reaction is most likely initiated by the activation of the 2′ hydroxyl on the ribose located at the scissile bond, and it results in a product with a 2′,3′-cyclic phosphate. The nature of the activation is still poorly understood, but it most likely involves a coordinated metal hydroxide. Notice in both (a) and (b) that the net number of bonds is conserved throughout the reaction.
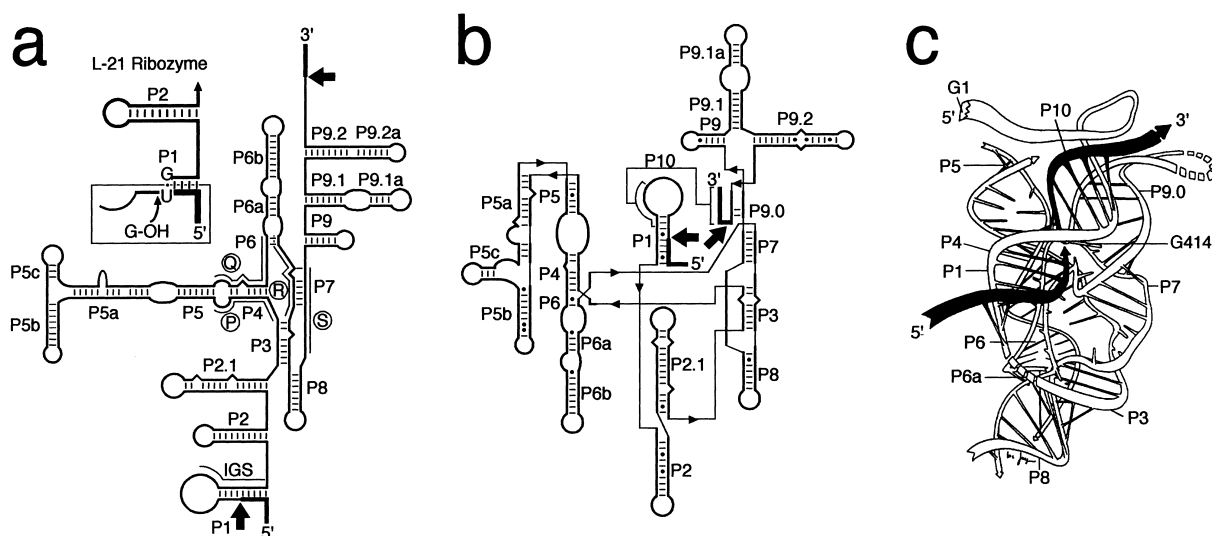
Fig. 2. The structure of group I introns. a: The 'classical' depiction of the secondary structure of the rRNA intron from *T. thermophila*, showing the phylogenetically conserved sequences P, Q, R, and S (circled) and the conserved pairings P1–P9. Nonconserved elements are designated by additional numbers or letters (e.g., P9.1 and P5c). The IGS sequence, which forms P1 and P10, is as indicated. The exon sequences are indicated as heavy lines and the splice sites by the arrows. The insert shows the 5′ end of the L-21 ribozyme, which is used in various reactions *in trans*, with a bound substrate. b: The new secondary structure depiction that more accurately represents the relationships of the different structural elements. The heavy lines are exon sequences and the very light ones are used to bridge the different elements, which are very close in tertiary space. c: A computer-generated model of the catalytic core of the *Tetrahymena* intron. The exon sequences are shown in solid and the intron as outline. This figure shows the intron after the first reaction, where the nonencoded guanosine, G1, is attached to the 5′ end of the intron and the 3′ hydroxyl of the 5′ exon is posed to make a nucleophilic attack at the 3′ splice site. This figure is reprinted with permission from [27].

sequences. These introns must be removed, at the RNA level, in order for the gene to be expressed functionally. There are five major categories of introns and splicing mechanisms. These consist of nuclear tRNA introns (reviewed by [18]), archaeal introns (reviewed by [19]), nuclear mRNA introns (reviewed by [20]), and the group I and group II introns (reviewed by [10,21]). Of these introns, some members of the group I and group II are clearly capable of catalyzing their own excision, in vitro, in an RNA-catalyzed fashion. However, genetic analyses have revealed protein factors that are often essential for group I and group II splicing in vivo [22]. The nuclear tRNA and archaeal introns clearly require protein factors (endonuclease and ligase). Nuclear mRNA introns may represent an intermediate state. These introns are spliced within large (40–60S) ribonucleoprotein particles (RNP), which consist of a number of small nuclear RNAs (snRNAs) and proteins [20]. However, it is increasingly thought that it is the snRNAs that are the

chemical engines of the complex as well as the determinants of the splice sites [23].

One of the oddities of group I and group II introns, which makes the term 'intron' something of an oxymoron, is that they sometimes encode proteins that are required in vivo either for splicing activity or for intron mobility (reviewed by [22]). The former are intron maturases, which can be highly specific for the intron that encodes it. The latter are endonucleases with reverse-transcriptase properties. These proteins are involved in intron 'homing', where crosses in yeast (the best studied genetically) result in the transfer of the intron into the intronless allele of the gene. There is also intron transposition (or 'retrohoming'), where the intron is inserted into other alleles. The widespread, but scattered, distribution of introns – especially introns with high sequence similarity occupying the same location in different organisms – suggests the possibility for horizontal transfer of introns; thus these introns may act somewhat akin to infectious agents. In group I introns,

mobility is apparently catalyzed by proteins alone and a functional (active) intron is not required. In group II, the intron itself is essential for activity and the intron RNA actually becomes inserted into the double-stranded DNA substrate [24,25].

3.1. Group I introns

Group I introns range in size from a few hundred nucleotides to around 3000. They are abundant in fungal and plant mitochondria, but they are also found in nuclear rRNA genes, chloroplast DNA (ctDNA), bacteriophage, eukaryotic viruses, and in the tRNA of ctDNA and eubacteria. In short, they are widely found except in higher eukaryotes (i.e., vertebrates; reviewed by [10,21,26]). The various group I introns have little sequence similarity, but they are characterized by four short conserved sequence elements, called P, Q, R, and S (Fig. 2a). P can always partially base-pair with Q, and R can always partially base-pair with S. In addition, all group I introns can fold into distinctive, phylogenetically conserved, secondary structures consisting of 10 paired segments (P1–P10; Fig. 2). Additional sequences, including large open reading frames (ORF), and structures are often found, but they do not disrupt the catalytic core, which consists of P3, P4, P6 and P7. However, their presence is used to subdivide the introns into various subgroups [27,28].

The secondary structure was originally deduced from computer modeling based on phylogenetic comparisons and, in the case of the *Tetrahymena* intron, on limited structural probing [29–31]. Additional evidence for this structure comes from enzymatic and chemical probing studies, additional phylogenetic analyses and from physical approaches (reviewed by [10,21,26,28]). These studies have confirmed the overall structure and revealed other important interactions, which have resulted in a new secondary structure depiction that more accurately reflects the spatial relationship of the different elements (Fig. 2b) [32].

The exon sequences are oriented relative to the catalytic core by base pairings with an intron sequence called the internal guide sequence (IGS), which constitutes P1 and P10 in the structure [29]. An additional pairing, called P9.0, further aligns the intron-3′ exon splice site [33,34]. It was so named
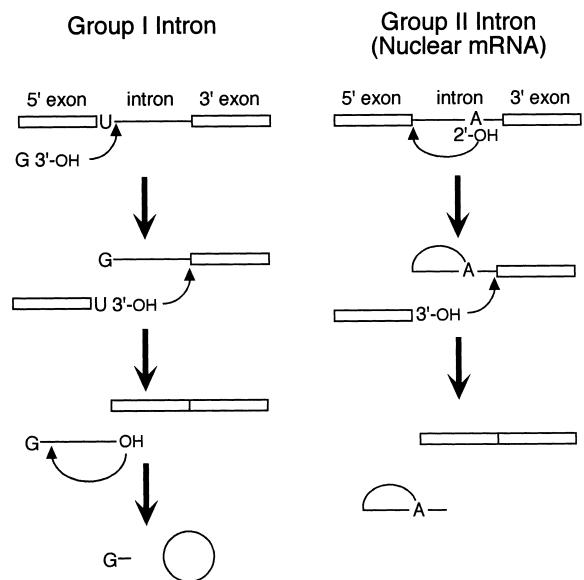

Fig. 3. Splicing mechanisms of group I and group II introns. In both cases, a series of *trans*-esterification reactions are used to excise the intron and ligate the exons. The net number of bonds remains the same throughout. The reaction is initiated by a guanosine cofactor in group I introns and by an internal adenosine in group II. The splicing reaction of nuclear pre-mRNAs follows the same pathway as group II introns, but it occurs on a large ribonucleoprotein complex. Details are described in the text.

because it occurs between P8 and P9 in the secondary structure, the latter being named before the interaction was discovered. The 5′ exon contains a highly conserved uridine on its 3′ end, which forms a functionally important U•G base pair with the IGS, and the intron contains a conserved guanosine on its 3′ end. Finally, there is a specific binding site for the guanosine cofactor that initiates the reaction. This site involves a conserved G-C base pair in P7 [35].

Michel and Westhof [27] have made a very clever, computer-generated, three-dimensional model of the catalytic core based on careful phylogenetic comparisons (Fig. 2c). This model brings all the relevant elements into proximity, and it has provided a sound basis for additional studies. A crystal structure was solved to 2.8 Å for a 160-nucleotide fragment of the *Tetrahymena* intron, consisting of P4–P6 [36]. While this structure contains less than half of the catalytic core of the enzyme and it tells us little about the reaction mechanism, it has revealed several exciting features. These include a tetraloop, GNRA (N is any

base and R is purine), docking with its receptor, several adenosine platforms, which are derived from adjacent adenosines forming a pseudo-base pair within the helix, and a ribose zipper, which involves a network of hydrogen bonds between the 2′-OH groups of ribose and acceptor groups of bases within the shallow groove of helices [36,37].

Recently, a 5.0-Å crystal structure has been solved for a 247-nucleotide-long fragment of the *Tetrahymena* intron, consisting of helices P3–P9 [38]. This new structure shows the previously characterized P4–P6 domain (helices P4, P5, and P6) largely unchanged and the P3–P9 domain (helices P3, P7, P8 and P9) wrapped around it. The close packing of the two domains creates a shallow cleft into which the P1 helix, containing the 5′ splice site, could fit. The structure also creates a particularly tight binding site in the P7 helix for the guanosine cofactor that initiates the splicing reaction. Unfortunately, the resolution is too low to visualize details within the structure, and the crystal structure is still missing the P1–P2 domain containing the splice site, so little can be determined about the reaction mechanism. However, it is clear that the computer generated model for the catalytic core [27,28] is largely consistent with this crystal structure. A triumph for computer modeling of RNA structures!

The splicing reaction of group I introns was first worked out for *Tetrahymena thermophila* (reviewed by [10,21,26]), and it will form the basis for my further discussion. Nevertheless, all characterized group I introns follow essentially the same pathway. The *Tetrahymena* intron is excised from the precursor rRNA by a two-step transesterification reaction (Fig. 3). The reaction is initiated by the nucleophilic attack of the 3′ hydroxyl of a guanosine cofactor at the 5′ splice site. The exon-intron phosphodiester bond is cleaved and the guanosine forms a 3′,5′-phosphodiester bond at the 5′ end of the intron. The now free 3′ hydroxyl of the 5′ exon then makes a nucleophilic attack at the 3′ splice site to form the ligated exons and release the intron with the non-encoded guanosine.

Nucleotides close to the 5′ end of the intron are realigned on the IGS and the highly conserved 3′ terminal guanosine of the intron makes a nucleophilic attack at a phosphodiester bond between nucleotides 15 and 16 or between nucleotides 19 and 20 within the intron in a reaction that is analogous to the first step of splicing. The intron is circularized (C-15 or C-19) and a small fragment containing the nonencoded guanosine is released. The circular product is also found in vivo, and its formation presumably helps to drive the reaction to completion. Although the circularization reaction is not universally conserved among group I introns, it is very common. Each step is essentially the forward or reverse of the same reaction, and the total number of phosphodiester bonds is conserved. Each step is fully reversible, and no external energy source is needed. The intron can fully reintegrate into rRNAs both in vitro and in vivo ([39] and reference therein). The phosphodiester bonds are inverted during the reaction, which is consistent with an S_N2, in-line, reaction mechanism (see [10,26]).

The circular product is normally considered the end product of the reaction, but the released intron retains catalytic activity and substrate specificity. For *T. thermophila*, a linear form of the intron lacking the first 21 and the last five nucleotides (L-21 *Sca*I) will catalyze a wide range of reactions on substrates added *in trans*. These include sequence-specific endonuclease, nucleotidyltransferase, ligase and phosphatase activities ([10,21,26]). Substrate specificity is changed by altering the IGS sequence. Because the reactions are catalyzed using the 3′ hydroxyl, they will work on both RNA and DNA substrates, although the latter has a much lower binding affinity and hence less activity. Since the L-21 *Sca*I RNA is unaltered in the reaction, it is considered a true RNA-based enzyme.

The properties of group I introns lend themselves to a number of different applications. The L-21 *Sca*I version of the *Tetrahymena* intron has been commercially sold as an RNA restriction enzyme. Circularly permutated precursor RNA, which contains end-to-end fused exons inserted within the middle of the intron sequence, is used to generate circular exon molecules by an 'inverse' splicing reaction [40]. A variant of this technique is used to generate circular, *trans*-cleaving, HDV ribozymes in vitro [41] and in vivo [42], which are more resistant to nucleases. A *trans*-splicing reaction can repair a truncated lacZ transcript in *Escherichia coli* [43] and in the cytoplasm of mammalian cells [44]. A *trans*-splicing intron can also change mutant β-globin transcripts, in
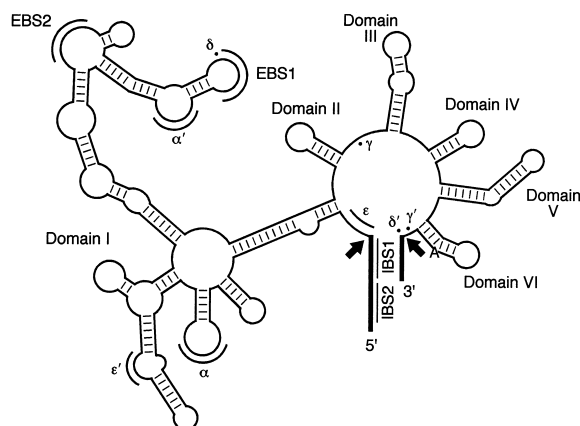
Fig. 4. The secondary structure of group II introns. This cartoon is a generalization from a number of different introns. The exon sequences are indicated as heavy lines and the splice sites by the arrows. The characteristics of the central wheel with the radiating domains is conserved, but the characteristics of the individual domains vary considerably. Tertiary interactions are formed between IBS1-EBS1, IBS2-EBS2, α-α′, ε-ε′, δ-δ′ and γ-γ′; the connections between these elements are not shown for clarity. The conserved adenosine used to initiate the splicing reaction is indicated in domain VI. Additional tertiary interactions have been identified, which vary with the intron (not shown).

sickle cell anemia, into mRNAs coding for antisickling γ-globin in human erythroid lineage cells [45].

3.2. Group II introns

Group II introns range in size from several hundred to around 2500 nucleotides. Although they are much less widely distributed than group I introns, they are found in fungal and plant mitochondria, in chloroplasts of plants, in algae, in eubacteria and especially in the chloroplasts of the protist *Euglena gracilis* (reviewed by [10,21,46,47]). Most are present in mRNAs, but a few also occur in tRNA and rRNA genes. In contrast to the extensively analyzed group I introns, much less is known about group II introns. This is partly because of their more limited distribution but also because very few of them are found to be self-splicing in vitro. Those that are autocatalytic require reaction conditions that are far from physiological (e.g., 100 mM $MgCl_2$, 500 mM $(NH_4)_2SO_4$ and 45°C; [48]).

The secondary structure of group II introns was originally deduced from phylogenetic comparisons and computer modeling [30,49]. It is normally de-

picted as six helical domains (I–VI) radiating as spokes from a central wheel (Fig. 4). On the basis of the structural features, group II introns are divided into two major subclasses, although some introns fit into neither class. Some contain long ORFs. Important structural features have been difficult to elucidate because of their poor reactivity in vitro and the difficulty of working with organelle-specific molecules. However, it appears that only domains I and V are indispensable [50]. In addition, domain V contains most of the relatively few phylogenetically conserved nucleotides found in group II introns and it might constitute the reaction center [46,47,51]. Domain VI contains the highly conserved adenosine that is generally used to initiate the splicing reaction. Except for domains I and V, the domains can be modified or deleted and the intron will retain some catalytic activity. Optional ORFs are frequently located in the loop of domain IV.

Degenerated forms of group II introns are found in plant chloroplasts and mitochondria that often lack recognizable cognates of the various domains. These are called group III introns (reviewed by [52]); they may require factors *in trans* for activity or be assembled from parts of group II introns. These degenerate introns provide a feasible pathway between group II and nuclear mRNA splicing, where more and more of the role of the intron is supplanted by *trans*-acting factors [23,47]. This possibility is made more plausible by the observation that some group II and group III introns occur within other group II or group III introns (called twintrons; [52]) and that others are discontinuous. Segments of these latter introns are transcribed within two or even three separate molecules from distant regions of the genome and the exons are assembled by a *trans*-splicing reaction (reviewed by [53]).

The 5′ exon is aligned by interactions between two intron binding sequences (IBS1 and IBS2), located near the 3′ end of the 5′ exon, and two exon binding sequences (EBS1 and EBS2), which are located in domain I of the intron (Fig. 4). The 5′ splice site is further defined by ε-ε′ interactions. These interactions are important both before and after 5′ cleavage. The 3′ splice site has multiple determinants, comprising δ-δ′, γ-γ′, and other as yet unidentified interactions. As yet, no tertiary model has been proposed.
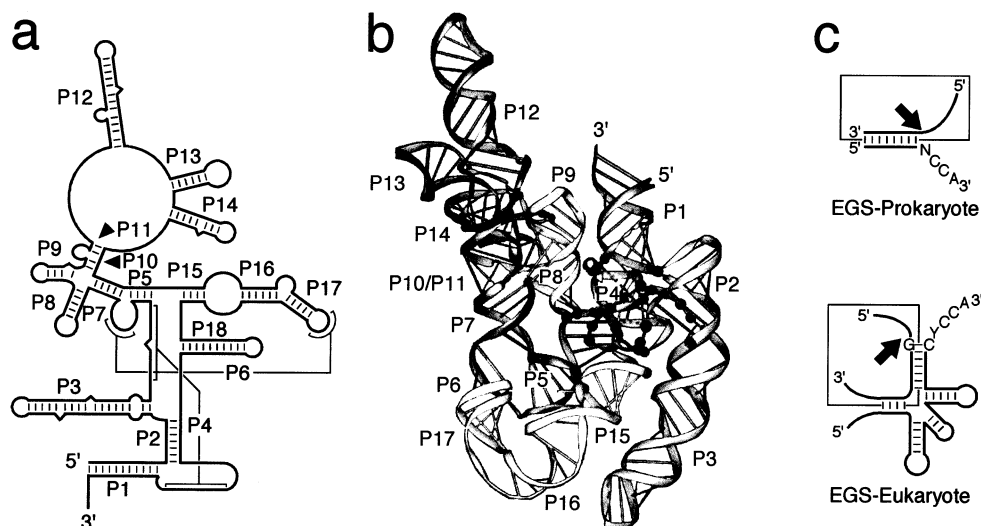
Fig. 5. Characteristics of RNase P RNA. a: The proposed secondary structure of M1 RNA, the RNA component of RNase P from *E. coli*. b: The computer-modeled tertiary structure of (a). The black spheres represent invariant nucleotides at the catalytic site and conserved nucleotides in the T-loop recognition site. Note that these conserved residues are clustered close to each other in what forms the catalytic core of the molecule. This figure is modified from [71]. c: Synthetic substrates for RNase P from prokaryotes and eukaryotes. The EGS, which is used to define the target specificity, is shown bound to the substrate (boxed) and the cleavage site is indicated by an arrow.

Unlike group I introns, where the first and second steps are thought of as forward and reverse steps of the same reaction, implying a single catalytic site, group II introns use two different nucleophiles (2′ and 3′ hydroxyls) and both steps show the same stereospecificity ([46,47] and references therein). This suggests that there are two independent reaction centers or that there is a single site that switches substrates [16].

The splicing reaction is generally initiated by the nucleophilic attack of the 2′ hydroxyl of a highly conserved adenosine in domain VI to form a distinctive structure, called a lariat, containing 3′-5′ and 2′-5′ phosphodiester bonds at the adenosine branch site (Fig. 3). The now free 3′ hydroxyl of the 5′ exon then makes a nucleophilic attack at the 3′ splice site to form the ligated exons and release the intron, still as a lariat. The reaction can also be initiated, both in vitro and in vivo, by the nucleophilic attack of water [48,54], although this is not a typical route. In this case, the intron is released as a linear molecule. As mentioned above, these reactions can also be catalyzed *in trans* from separate transcripts. The splicing mechanism and the lariat product are reminiscent of those found in nuclear mRNA splicing,

and it provides an additional evolutionary link between the two splicing reactions [23,47].

The applications of group II introns are more limited than those of group I. Nevertheless, it can catalyze the cleavage of ligated precursors using the energy of a phosphoanhydride bond, ligate RNA to DNA, and cleave single-stranded DNA substrates [55]. It is also used to circularize exon sequences [56], as described above for group I introns. However, the fact that the intron RNA becomes inserted into double-stranded DNA during intron homing [24,25] suggests that other applications will soon be found.

4. RNase P

RNase P is a ubiquitous enzyme in all characterized organisms that processes the 5′ termini of tRNA precursors (reviewed by [10,57–60]). In eubacteria, RNase P exists as a ribonucleoprotein complex, consisting of a large RNA of about 350–400 nucleotides (M1 RNA in *E. coli*) and a small basic protein of ∼14 kDa (C5 in *E. coli*). The basic protein is essential for activity in vivo, however the RNase P RNA,

by itself, can catalyze the reaction in vitro [2]. This reaction requires high salt concentrations (e.g., 1 M K^+ or NH_4^+ and 10 mM $MgCl_2$; [60]), which suggests that the basic protein component acts only as an electrostatic shield to promote binding between the RNA enzyme and the RNA substrate. However, the protein also affects cleavage-site specificity and turnover, so its full role in the reaction is still unclear (reviewed by [58,60]). RNase P from *E. coli* will also process other substrates that partially resemble tRNAs (e.g., 4.5S RNA [61]), but these reactions appear to be relatively minor. RNase P RNA is the only characterized ribozyme that, unmodified, acts *in trans* on multiple substrates, and hence it is considered the only true, naturally occurring, RNA enzyme.

RNase P from the nuclei and mitochondria of eukaryotes also exist as ribonucleoprotein complexes, although they generally have much higher protein contents (ca. 50–70% versus ca. 10% for eubacteria [57]). Moreover, while the RNA component is inevitably found to be essential, the RNase P RNA from eukaryotes has never been shown to have catalytic activity. Archaebacteria have many properties that make them more similar to the eukaryotes than to the eubacteria, and they likewise are not known to have RNA-alone catalytic activity [60].

There is evidence that another ribonuclease in eukaryotes, RNase MRP, is closely related, and perhaps homologous (i.e., evolutionarily related), to RNase P (reviewed by [62,63]). It is a ribonucleoprotein complex that participates in nucleolar pre-rRNA processing. The RNA component can fold into similar secondary structures as RNase P RNA and, in yeast, it shares common protein components [64]. However, like the eukaryotic RNase P, the RNase MRP RNA has not been shown to have catalytic activity by itself.

Although there is little sequence conservation, all the eubacterial RNase P RNAs can be folded into similar, although not identical, secondary structures on the basis of comparative sequence analyses [60,65,66]. The *E. coli* M1 RNA consists of 18 paired helices (P1–P18; Fig. 5a). The *Bacillus subtilis* M1 RNA analog is more diversified from *E. coli* than many of the other eubacteria. It folds into a similar structure, but P6, P13, P14, P16 and P17 are missing and it contains extra helices P5.1, P10.1, P15.1, and

P19 (not shown). These two RNAs are the most extensively studied, but through a careful comparative analysis of the different eubacteria, it is possible to derive a common core structure consisting of helices P1–P5, P7–P12 and P15 [60]. However, the activity of this 'minimal' structure has not yet been demonstrated.

Native RNase P RNAs differ from the phylogenetic minimum by having extra stems and stem-loop structures. While probably not needed for catalysis, these structures nevertheless lower the ionic strength requirements and enhance their thermal stability [60]. Such elements may be redundant in the sense that any single one can be deleted or modified without significantly altering the activity, but it is not possible to simultaneously delete or modify all of them. The RNase P RNA from eukaryotes and archaebacteria have little sequence similarity to their eubacterial counterparts, but they can often be folded into a universally conserved core structure as well [67]. Missing RNA elements from the structures may be supplanted by the protein component(s), but this has not been characterized.

Three-dimensional models of the *E. coli* M1 RNA have been made by Westhof and Altman [68] and by Harris et al. [69], and both groups have recently refined these models [70,71] (Fig. 5b). Three-dimensional models for the *B. subtilis* RNase P RNA are also available [70,71]. These are computer-derived models based on data obtained from phylogenetic comparisons, mutational analyses, chemical probing and from crosslinking studies. It is beyond the scope of this review to discuss these models, especially in light of their speculative nature, but generally the models have a good fit to the experimental data. Both groups' models have similar overall structures and both provide a pocket or cleft into which the tRNA substrate will fit. However, many of the specific interactions vary; thus these models are expected to undergo continual refinement as additional experimental data become available.

RNase P uses water as a nucleophile to cleave the phosphodiester bond (Fig. 1a). The exact mechanism by which the tRNA precursor is bound is still unclear. The 3′ half of the acceptor stem is thought to function as an external guide sequence (EGS), but it does not uniquely define the cleavage site. The primary sequence does not seem to be important nor

does any single element uniquely define the cleavage site. Instead, recognition could result from several redundant factors, including the distance along the coaxially stacked T stem-loop and acceptor stem, the 3′ terminal CCA sequence, the EGS alignment, and a conserved guanosine 3′ to the cleavage site (reviewed by [58,59]). Thus, recognition is largely, if not entirely, based on tertiary interactions with the substrate.

From studies using small substrates, it is possible to design synthetic EGSs that can target any RNAs for cleavage by RNase P in vitro or in vivo (Fig. 5c; reviewed by [72]). The basis for this is that the EGS binds to the target RNA and makes it look like a tRNA substrate. In eubacteria, the minimal requirement is that it forms a short stem with a free NCCA (Fig. 5c). The target sequence can be virtually anything. In eukaryotes, the EGS-substrate complex must more closely resemble a tRNA (Fig. 5c). This reaction works in vitro, in bacteria and in human cells [72]. A variant of this technique is to add the appropriate EGS (here called an internal guide sequence or IGS) to the 3′ end of the RNase P RNA. This increases the efficiency of the reaction, and it is used to inactivate thymidine kinase mRNA from herpes simplex virus in cell lines [73]. Yet, while the therapeutic potential of RNase P has been demonstrated, it has not been widely used in therapeutic applications.

5. General properties of small catalytic RNAs

Self-cleaving RNAs are generally found in small ($\sim$220 to $\sim$460 nucleotides long) RNA pathogens of plants known as viroids, virusoids and linear satellite viruses (reviewed by [13,74]). However, they are also found within satellite RNAs of salamanders, *Neurospora*, and within another pathogenic satellite virus found in man. The viroid and satellite RNAs are generally replicated by an RNA-dependent rolling-circle mechanism, and the catalytic domains are thought to process the linear concatemers that are generated into unit-length progeny.

The linear, unit-length progenies produced during replication in vivo are subsequently ligated to form closed-circular molecules that are used in the next round of rolling-circle replication. It is reasonable to expect that this is catalyzed by the ribozyme as well since, mechanistically, it represents the reverse of the cleavage reaction, and it would be analogous to the splicing reaction carried out by the group I and group II introns. However, in vitro only the hairpin ribozyme shows significant ligation activity. Protein factors may be involved in vivo or other, as yet unidentified, RNA elements may be required.

Four motifs are characterized (described below), and they all catalyze reactions that generate products with 2′,3′-cyclic phosphates and 5′ hydroxyls (reviewed by [13,74]; Fig. 1b). As with the previous ribozymes, they all require a divalent metal ion, normally Mg^{2+}, for activity.

The intramolecular self-cleaving activity is converted into a *trans*-cleaving activity by making the 'substrate' and 'ribozyme' into separate molecules. However, the 'substrate' remains an integral part of the structure of the active ribozyme, and hence the ribozyme is often simply defined as the unmodified portion of the molecule and the substrate is the cleaved portion. Hence, the 'ribozyme' may consist of different sequence elements, depending on the construct. The catalytic domains of these ribozymes are small and relatively well characterized, and they are more widely used in therapeutic applications. Each has characteristics that confer specific advantages and disadvantages as therapeutic agents. The in vitro and ex vivo activity of *cis*-cleaving forms of three of these self-cleaving ribozymes (hammerhead, hairpin and HDV) have been compared [75].

5.1. Hammerhead ribozyme

The hammerhead ribozyme is probably the most extensively studied of all the ribozymes, and it is the motif most commonly found in the viroids and satellite RNAs (reviewed by [13]). Currently 16 hammerhead motifs are known in the plus and minus strands of these plant pathogens. Three other hammerhead motifs are found in the satellite 2 RNAs from the salamanders, *Triturus vulgaris*, *Ambystoma talpoideum* and *Amphiuma tridactylum* ([13] and references therein). This ribozyme was so named because its Australian discoverers found the secondary structure, as originally drawn, to be reminiscent of the head of a hammerhead shark. It is the smallest of
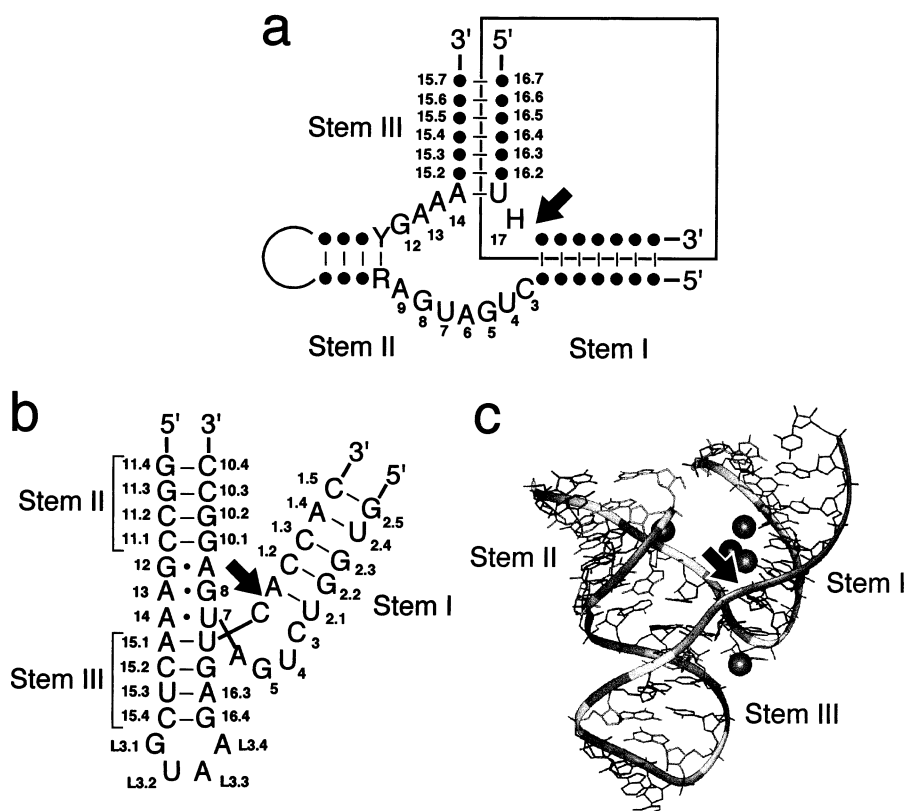
Fig. 6. Characteristics of the hammerhead ribozyme. a: The secondary structure of the hammerhead ribozyme showing the conserved sequence and structure. The dots represent nucleotides that can be anything, Y is a pyrimidine, R is a purine and H is any nucleotide except guanosine. The arrow indicates the self-cleavage site. The boxed region shows the portion that is normally the substrate in *trans*-cleaving versions of the ribozyme. The numbering is based on standardized nomenclature [104]. b: A new secondary structure drawing that more accurately reflects the spatial relationships of the different elements. This structure differs from (a) in that it shows a loop in stem III rather than in stem II. c: The solved crystal structure of the sequence shown in (b). This is an RNA-only structure infused with Mg^{2+}, shown as spheres. Only the Mg^{2+} ion close to the scissile bond is generally accepted as being functionally relevant (see text). This figure is modified from [81].

the naturally occurring self-cleaving RNAs, at 40–50 nucleotides in length.

Analyses of the hammerhead ribozyme are voluminous (reviewed by [9,76–78]). It consists of three helical regions, which are variable, and three single-stranded regions that contain most of the highly conserved nucleotides (Fig. 6a). The length of the helical arms can be quite variable, and helix II can be reduced to two base pairs. Mutating any of the conserved residues markedly reduces activity; consequently, important functional groups are often identified by incorporating synthetic nucleotide analogs into the RNA (e.g., inosine for guanine; reviewed by [76]).

Cleavage occurs after an NUH triplet, where N is any nucleotide, and H is any nucleotide except guanosine. The most effective triplet is GUC, but other triplet combinations will work nearly as well; their relative activities have been compared, although the ordering can vary depending on the method of analysis (reviewed by [8,9,77]). The reaction mechanism is extensively studied, and it appears to involve a metal-coordinated hydroxide, which probably directly activates the 2′ hydroxyl ([9,76,78] and reference therein). The reaction products are consistent with an S_N2 (in-line) reaction mechanism; this was suggested by an inversion of the phosphate at the scissile linkage.
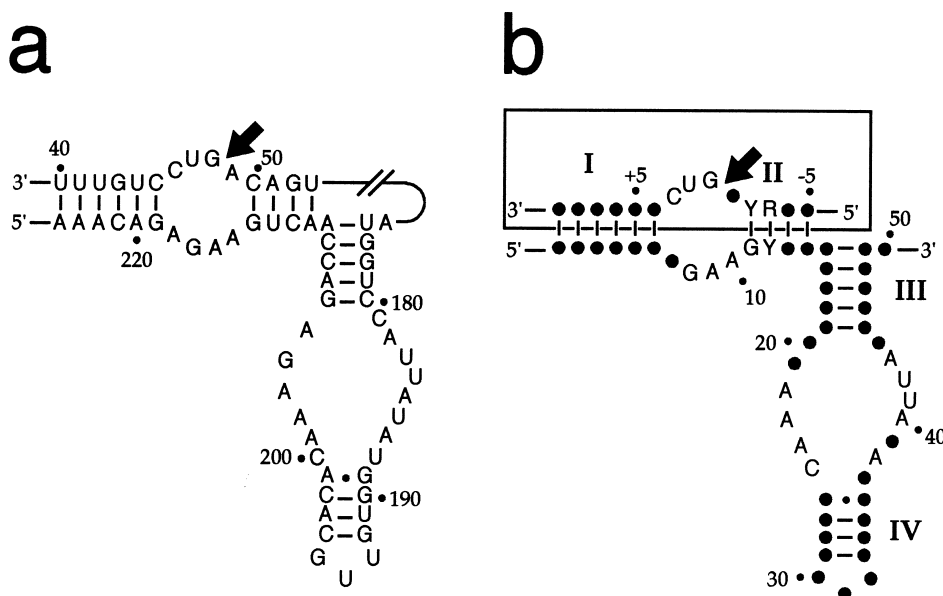
Fig. 7. Secondary structure of the hairpin ribozyme. a: The minus strand of sTRSV (numbering is that of the full-length virus). The arrow shows the cleavage site. b: The consensus sequence and structure, where dots are any nucleotide, Y is a pyrimidine and R is a purine. The boxed region represents the portion that is normally the substrate in *trans*-cleaving reactions. The substrate is numbered relative to the cleavage site and the 'ribozyme' relative to the 5′ end. This figure is modified from [83] and it incorporates recent experimental data from [105].

The hammerhead ribozyme was the first catalytic RNA for which the complete X-ray crystal structure was solved. There are now a number of such structures available with resolutions ranging from 2.6 Å to 3.1 Å, as well as a structure derived from fluorescence resonance energy transfer (FRET; reviewed by [76,78,79]). The molecule has a Y shape (Fig. 6b,c), with stem I and stem II at the arms and stem III at the base. Recently, the crystallized structure of a reaction intermediate was determined using a tallo-5′-C-methyl-ribose-modified ribozyme that is kinetically blocked for the final cleavage reaction [80]. This latter structure is more compatible with an S_N2 reaction mechanism. This is in contrast to the previously solved structures that showed the ribozyme in a ground state that was incompatible with such a mechanism. A second Mg^{2+} binding site has potentially been identified in another crystal structure that could be involved in stabilizing the pentacoordinated phosphate transition state [81]. However, the existence of this site is still debated (see [78]). Moreover, there are other incompatibilities between the exper-imental data and the crystalline structures [78], and clearly additional work will be needed.

The hammerhead ribozyme is divided into separate 'ribozyme' and 'substrate' in several ways, but the one shown in Fig. 6a is the most commonly used because most of the conserved residues are contained within the ribozyme rather than in the substrate. The hybridizing arms are varied to optimize ribozyme activity and substrate specificity. Normally hybridizing arms of six or seven base pairs are considered optimal, but for variable arms it is better to have a long stem III and short stem I than the reverse (reviewed by [9]). Often, a tetraloop (frequently GAAA) is used for loop II and stem II is GC-rich to further increase the stability of the stem-loop. The malleability of the hammerhead ribozyme makes it the most commonly used ribozyme for in vivo studies, and there are many successful examples (reviewed by [7–9]). A *trans*-cleaving hammerhead ribozyme is approved for phase II clinical trials against HIV-1 by Ribozyme Pharmaceuticals (RPI).

5.2. Hairpin ribozyme

The hairpin ribozyme is found in three pathogenic, plant, satellite viruses, although the one found in the satellite virus associated with tobacco ring spot virus (sTRSV) is the best characterized. It consists of four stem regions that, when lined up coaxially, somewhat resemble a hairpin; interestingly, it was also originally named 'paperclip,' which may, in fact, better represent its overall three-dimensional shape (see [13]; Fig. 7a). It consists of two noncontiguous sequences of 50 and 14 nucleotides within the minus strand of sTRSV (reviewed by [13,17,82–84]. The secondary structure was determined based on computer-aided modeling, limited phylogenetic comparisons, mutational analyses and by in vitro selection.

The other hairpin ribozymes are found in the satellite viruses of arabis mosaic virus (sARMV) and chicory yellow mottle virus (sCYMV), and they mostly differ from the sequence shown in Fig. 7a by nucleotide changes within the helical regions that maintain the structure as shown. Indeed, the mutational and in vitro selection analyses show that the helical regions are structural elements that can largely be changed, as long as the integrity of the helices is maintained. Most of the conserved nucleotides occur within the single-stranded regions. The guanosine 3′ to the cleavage site is essential, but altering the other conserved positions can dramatically reduce the activity as well. The bulged region between helices III and IV contain a conserved motif, called a UV-loop motif, that is found in a diverse group of RNAs, including viroids, 5S rRNA and the sarcin-ricin loop of 28S rRNA [82]. However, the role this motif plays in catalysis is still unknown. The consensus sequence and structure are shown in Fig. 7b.

Recently, a computer-generated tertiary model was made that was based on preexisting structural data and on the spatial distance of tolerated, interdomain, aryl-disulfide crosslinks [85]. Additional information was also obtained by Walker et al. [86] using FRET data. The current model shows helix I coaxially stacked on helix II and helix IV coaxially stacked on helix III. These two extended helices are then bent so that helix II and helix III, and helix I and helix IV are positioned side by side. This places the two highly conserved bulged regions in proxim-

ity, and they could thus form the catalytic core. However, this tertiary model is still preliminary and additional data are required before the details of the catalytic site are known.

A major advantage of the hairpin ribozyme lies in its ability to catalyze both cleavage and ligation reactions efficiently in vitro; this has greatly facilitated in vitro selection experiments because new substrates, with the appropriate PCR primer sites, are easily generated ([82] and references therein). The RNA-catalyzed ligation reaction is also thought to be relevant in vivo, in that the RNA can both cleave the linear multimers generated during rolling-circle replication and ligate them to form the circular RNA progeny. However, as with the other catalytic motifs used in viroid replication, the cleavage-ligation reaction must be carefully regulated in vivo to prevent inappropriate cleavage, or ligation, of the resulting progeny. The mechanism by which this is accomplished is still unknown. Like the other catalytic RNAs, the hairpin ribozyme reaction requires a divalent cation. However, it appears that, unlike the hammerhead ribozyme, the hydrated cation (usually Mg^{2+}) is not directly coordinated to the phosphate; moreover, the essential role of divalent cations has recently been called into question (see [17]).

Since the 'substrate' portion of the *cis*-cleaving hairpin ribozyme is discontinuous with the rest of the molecule, it is obvious where to separate the two domains (boxed region in Fig. 7b). The substrate should contain the sequence RYN*GUC, where R is a purine, Y is a pyrimidine and N is any nucleotide. Cleavage occurs at the position indicated by a *. Stem II should be four base pairs, but stem I can be significantly extended (e.g., [87]), as can stem IV [88]. Substrate specificity is changed by altering the nonconserved residues within the base-paired region. The hairpin ribozyme is used to target HIV-1 RNA in cell culture, and it is currently approved for clinical trials (see [84]). However, its effective use as a *trans*-cleaving ribozyme in vivo is still rather limited; the reasons for this are unclear.

5.3. VS RNA ribozyme

The mitochondria of certain strains of *Neurospora* contain the Varkud plasmid (a retroplasmid), which encodes a reverse transcriptase, and a small, unre-
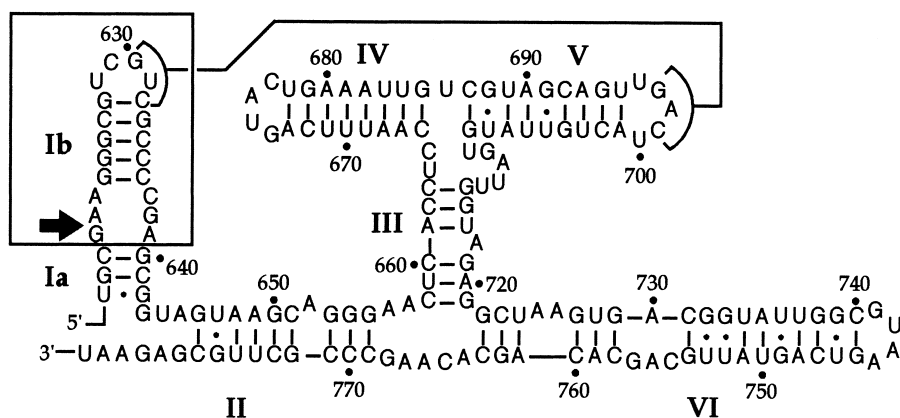
Fig. 8. Secondary structure of the VS RNA ribozyme. The arrow shows the cleavage site and numbering is that of the full-length VS RNA. The *trans*-cleaving form consists of nucleotides 640–881 and the substrate is shown boxed.

lated, RNA (VS RNA). The VS RNA is transcribed from circular or multimeric VS plasmid DNA by a mitochondrial RNA polymerase, and the resulting transcripts are subsequently site-specifically cleaved and ligated to form circular, 881 nucleotides long, RNA monomers [89]. These monomers are then reverse transcribed and made double stranded to form the mature VS plasmid.

In vitro transcribed VS RNA precursors are cleaved and ligated by the RNA itself and this is presumed to occur in vivo as well ([90] and references therein). Of all the self-cleaving RNAs, the catalytic properties of VS RNA are the most poorly understood. At 154 nucleotides long it is also the largest. The minimal sequence that retains catalytic activity contains one nucleotide 5′ and 153 nucleotides 3′ to the cleavage site. However, this structure can be reduced to 121–126 nucleotides by making internal deletions within the helices [91]. An RNA secondary structure is proposed, but except for a tertiary interaction between loop I and loop V, little is known about its overall conformation (Fig. 8).

The catalytic domain of VS RNA is converted into a *trans*-cleaving ribozyme by using a 144-nucleotide fragment of the VS RNA from 640 to 881 (VS RNA numbering; [92]). The minimal substrate consists of one nucleotide 5′ and 19 nucleotides 3′ to the cleavage site, and it forms a short stem-loop structure. As with RNase P, the ribozyme seems to recognize the structure of the substrate largely as a helical domain.

The minimal sequence requirement 5′ to the cleavage site is a characteristic shared only with the HDV ribozyme (see below) and it could make this ribozyme suitable for 3′ end trimming of RNAs expressed in vitro or in vivo. However, the uncertainty in the substrate requirements and the lackadaisical activity *in trans* have limited its application, although recent experiments have improved its activity [93].

5.4. HDV ribozyme

The hepatitis delta virus (HDV) is a viroid-like satellite virus of the hepatitis B virus (HBV), and it is the sole example of such a virus in mammalian systems (reviewed by [94,95]). It is widespread and can cause severe fulminant hepatitis in infected patients. It is about 1700 nucleotides long, and it encodes a single protein that is expressed in two forms due to an RNA editing event. Both the genomic, infectious strand, and the antigenomic strand have self-cleaving domains (reviewed by [96–99]). Despite previous pronouncements, no biologically relevant, RNA-catalyzed, ligation reaction has been observed in vitro, although the integrity of the RNA catalytic domains is clearly essential for both the cleavage and ligation reaction in vivo ([99] and references therein). A possible mechanism for the biological control of these reactions, to prevent inappropriate cleavage or ligation, has been proposed [96].

The minimal domain containing self-cleaving activity has one nucleotide 5′ and 84 nucleotides 3′ to
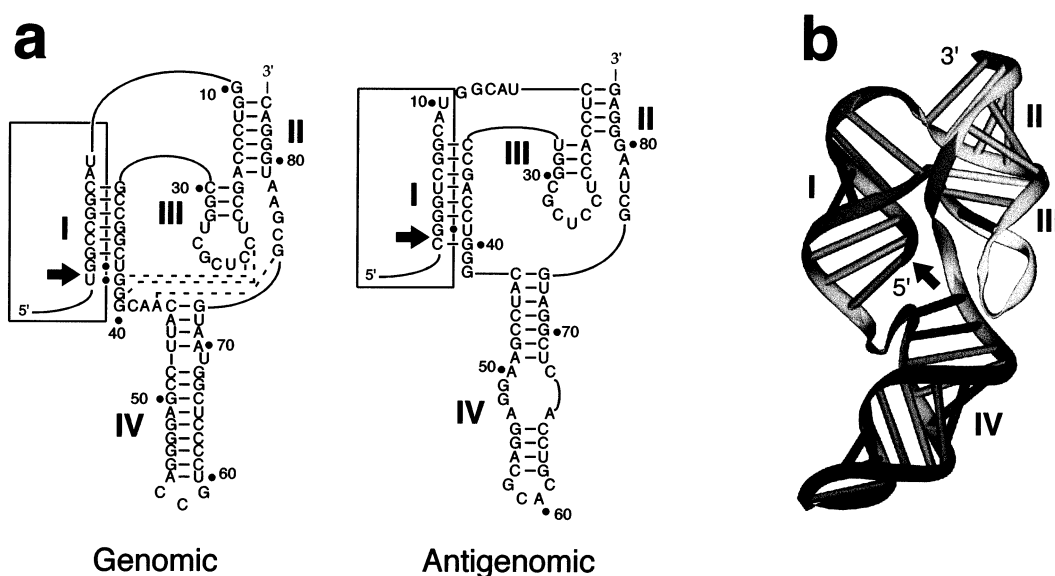
Fig. 9. Characteristics of the HDV ribozyme. a: The genomic and antigenomic ribozymes from HDV. Numbering of the nucleotides is relative to the cleavage sites, indicated by arrows. Helical domains are separated by lines to facilitate the presentation. The length of stem I is critical, although there is no evidence that the −1 base pair is needed. The typical *trans*-cleaving forms are separated in junction I/II, where the substrates are shown boxed. The identities of the nonbase-paired residues in the substrate (boxed) are not important for *trans*-cleaving activity. Additional base-pair interactions, that were recently derived for the genomic ribozyme, are shown as dashed lines. b: Computer-generated three-dimensional model of the genomic catalytic domain. This figure is modified from [100]. Recently, the crystal structure for the genomic HDV ribozyme was obtained [103]. This structure is similar to that shown here except that there is an additional pseudoknot interaction between C21 and C22 with G38 and G39. Moreover, G10 is stacked on helix II and A43 and G74 stack on hairpin IV.

the cleavage site for both domains. Despite the sequence differences, both sequences fold into similar secondary structures, of which the pseudoknotted structure shown in Fig. 9a is now the most widely accepted. It consists of four stem regions; three of these stems (I, II and IV) are largely structural elements, while the specific sequences in hairpin III and in the junctions I/IV and IV/II are more important. The catalytic domains of HDV are known for their ability to retain *cis*-cleaving activity at high temperatures and in the presence of denaturants (see [97] and references therein). The tertiary structures have been computer modeled for the genomic and antigenomic forms of the pseudoknot model ([100,101]; Fig. 9b) and for the antigenomic axehead variant [102].

Recently, a 2.3-Å crystal structure has been solved for the genomic HDV ribozyme [103]. This was accomplished by replacing hairpin IV with a small hairpin structure that binds tightly to the protein U1A, a spliceosomal protein. By co-crystallizing

the RNA with the protein, the authors were better able to obtain highly structured crystals that diffracted to a high resolution. It also greatly facilitated heavy metal substitution that is necessary for obtaining crystal phasing. This structure is very similar to the computer-predicted model, but it revealed some unexpected results. There is an additional pseudoknot structure derived from base pairs between C21 and C22, in loop III with G39 and G38, at the base of helix I. These specific interactions were not previously predicted, although the importance of the nucleotides were correctly derived [100] In addition, A43 and G74 stack on the end of helix IV to form noncanonical base pairs and G10 forms an extension to helix II as previously predicted [97,98]. These interactions create a structure where helix IV is rotated relative to the computer model and hairpin III is more compressed. Nevertheless, the two structures are otherwise very similar. The crystal structure provides an organized, almost protein-like, crevice for the active site. Unfortunately, it reveals little

about the reaction mechanism or the role of the Mg^{2+} ions.

As with the other ribozymes, the *cis*-cleaving activity of the HDV ribozymes can be converted into a *trans*-cleaving activity (reviewed by [97–99]). The most common form is indicated in Fig. 9a. However, other permutations are possible, including a large substrate consisting of sequences −5 to ∼60 (numbering in Fig. 9a) and a small ribozyme consisting of sequences from ∼60 to +84. Since most of the conserved elements are contained within the substrate, the practical utility of this latter form is somewhat limited. A completely closed-circular variant of the *trans*-cleaving ribozyme shown in Fig. 9a also has been generated; a short loop was used to close the end of stem II [41]. The absence of free ends makes this ribozyme particularly resistant to the exonucleases found in serum and the cellular environment. It is possible to change the substrate binding sequence to target other RNAs. In theory, most substrate sequences are possible, although a guanosine at the −1 position, relative to the cleavage site, is inhibitory and a purine-pyrimidine base pair at position +1 is preferred. However, in practice many of the changes in the substrate-binding sequence have unpredictable effects.

Up to now, the HDV ribozyme has a rather limited use in therapeutic applications, largely because of the difficulty of obtaining high activity with the *trans*-cleaving forms. However, because of the minimal sequence requirements 5′ to the cleavage site, the *cis*-cleaving activity is useful for generating discrete 3′ ends of RNA. This has applications, for example, in processing ribozyme cassettes (transcripts containing multiple ribozyme units), for generating homogeneous ends on in vitro transcribed RNA or for the expression of discrete viral RNA transcripts off plasmid DNA ([99] and references therein).

6. Concluding comments

In this review I have attempted to summarize briefly the important features of the different catalytic RNAs that have so far been identified. The ribozyme field has advanced far in an incredibly short time. There are seven ribozymes identified in nature, and all of them have been engineered to cleave or modify other RNAs *in trans*. Other ribozymes have been created de novo, and they can catalyze a variety of reactions. Moreover, other cellular processes have significant RNA components (e.g., the spliceosome and ribosomes) where RNA catalysis may play an important role. Additional ribozymes are bound to be discovered in the future. Finally, the disappointment that many earlier researchers had when working with ribozymes in vivo has now opened up to new opportunities as people have discovered new ways of dealing with the intracellular environment. Ribozymes are now an important component of future developments in gene regulation. They are worth keeping an eye on.

References

[1] Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E. and Cech, T.R. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. Cell 31, 147–157.

[2] Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell 35, 849–857.

[3] Noller, H.F., Hoffarth, V. and Zimniak, L. (1992) Unusual resistance of peptidyl transferase to protein extraction procedures. Science 256, 1416–1419.

[4] Nitta, I., Kamada, Y., Noda, H., Ueda, T. and Watanabe, K. (1998) Reconstitution of peptide bond formation with *Escherichia coli* 23S ribosomal RNA domains. Science 281, 666–669.

[5] Jaeger, L. (1997) The New World of ribozymes. Curr. Opin. Struct. Biol. 7, 324–335.

[6] Joyce, G.F. and Orgel, L.E. (1993) Prospects for understanding the origin of the RNA world. In: The RNA World (Gesteland, R.F. and Atkins, J.F., Eds.), pp. 1–25. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[7] Marschall, P., Thomson, J.B. and Eckstein, F. (1994) Inhibition of gene expression with ribozymes. Cell. Mol. Neurobiol. 14, 523–538.

[8] Kijima, H., Ishida, H., Ohkawa, T., Kashani-Sabet, M. and Scanlon, K.J. (1995) Therapeutic applications of ribozymes. Pharmacol. Ther. 68, 247–267.

[9] Birikh, K.R., Heaton, P.A. and Eckstein, F. (1997) The structure, function and application of the hammerhead ribozyme. Eur. J. Biochem. 245, 1–16.

[10] Cech, T.R. (1993) Structure and mechanism of the large catalytic RNAs: group I and group II introns and ribonuclease P. In: The RNA World (Gesteland, R.F. and Atkins, J.F., Eds.), pp. 239–269. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[11] Symons, R.H. (1994) Ribozymes. Curr. Opin. Struct. Biol. 4, 322–330.

[12] Scott, W.G. and Klug, A. (1996) Ribozymes: structure and mechanism in RNA catalysis. Trends Biochem. Sci. 21, 220–224.

[13] Symons, R.H. (1997) Plant pathogenic RNAs and RNA catalysis. Nucleic Acids Res. 25, 2683–2689.

[14] Tanner, N.K. (1998) Ribozymes: caractéristiques et applications. Virologie 2, 127–137.

[15] Narlikar, G.J. and Herschlag, D. (1997) Mechanistic aspects of enzymatic catalysis: lessons from comparison of RNA and protein enzymes. Annu. Rev. Biochem. 66, 19–59.

[16] Steitz, T.A. and Steitz, J.A. (1993) A general two-metal-ion mechanism for catalytic RNA. Proc. Natl. Acad. Sci. USA 90, 6498–6502.

[17] Walter, N.G. and Burke, J.M. (1998) The hairpin ribozyme: structure, assembly and catalysis. Curr. Opin. Chem. Biol. 2, 24–30. (correction Curr. Opin. Chem. Biol. 2, 303)

[18] Abelson, J., Trotta, C.R. and Li, H. (1998) tRNA splicing. J. Biol. Chem. 273, 12685–12688.

[19] Lykke-Andersen, J., Aagaard, C., Semionenkov, M. and Garrett, R.A. (1997) Archaeal introns: splicing, intercellular mobility and evolution. Trends Biochem. Sci. 22, 326–331.

[20] Moore, M.J., Query, C.C. and Sharp, P.A. (1993) Splicing of precursors to mRNA by the spliceosome. In: The RNA World (Gesteland, R.F. and Atkins, J.F., Eds.), pp. 303–357. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[21] Saldanha, R., Mohr, G., Belfort, M. and Lambowitz, A.M. (1993) Group I and group II introns. FASEB J. 7, 15–24.

[22] Lambowitz, A.M. and Belfort, M. (1993) Introns as mobile genetic elements. Annu. Rev. Biochem. 62, 587–622.

[23] Newman, A. (1997) RNA splicing: out of the loop. Curr. Biol. 7, R418–R420.

[24] Grivell, L.A. (1996) Transposition: Mobile introns get into line. Curr. Biol. 6, 48–51.

[25] Matsuura, M., Saldanha, R., Ma, H., Wank, H., Yang, J., Mohr, G., Cavanagh, S., Dunny, G.M., Belfort, M. and Lambowitz, A.M. (1997) A bacterial group II intron encoding reverse transcriptase, maturase, and DNA endonuclease activities: biochemical demonstration of maturase activity and insertion of new genetic information within the intron. Genes Dev. 11, 2910–2924.

[26] Cech, T.R. (1990) Self-splicing of group I introns. Annu. Rev. Biochem. 59, 543–568.

[27] Michel, F. and Westhof, E. (1990) Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. J. Mol. Biol. 216, 585–610.

[28] Lehnert, V., Jaeger, L., Michel, F. and Westhof, E. (1996) New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: A complete 3D model of the *Tetrahymena thermophila* ribozyme. Chem. Biol. 3, 993–1009.

[29] Davies, R.W., Waring, R.B., Ray, J.A., Brown, T.A. and Scazzocchio, C. (1982) Making the ends meet: a model for RNA splicing in fungal mitochondria. Nature 300, 719–724.

[30] Michel, F., Jacquier, A. and Dujon, B. (1982) Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure. Biochimie 64, 867–881.

[31] Cech, T.R., Tanner, N.K., Tinoco Jr., I., Weir, B.R., Zuker, M. and Perlman, P.S. (1983) Secondary structure of the *Tetrahymena* ribosomal RNA intervening sequence: Structural homology with fungal mitochondrial intervening sequences. Proc. Natl. Acad. Sci. USA 80, 3903–3907.

[32] Cech, T.R., Damberger, S.H. and Gutell, R.R. (1994) Representation of the secondary and tertiary structure of group I introns. Nature Struct. Biol. 1, 273–280.

[33] Burke, J.M., Esherick, J.S., Burfeind, W.R. and King, J.L. (1990) A 3′ splice site-binding sequence in the catalytic core of a group I intron. Nature 344, 80–82.

[34] Michel, F., Netter, P., Xu, M.-Q. and Shub, D.A. (1990) Mechanism of 3′ splice site selection by the catalytic core of the sunY intron of bacteriophage T4: the role of a novel base-pairing interaction in group I introns. Genes Dev. 4, 777–788.

[35] Michel, F., Hanna, M., Green, R., Bartel, D.P. and Szostak, J.W. (1989) The guanosine binding site of the *Tetrahymena* ribozyme. Nature 342, 391–395.

[36] Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Kundrot, C.E., Cech, T.R. and Doudna, J.A. (1996) Crystal structure of a group I ribozyme domain: principles of RNA packing. Science 273, 1678–1685.

[37] Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Szewczak, A.A., Kundrot, C.E., Cech, T.R. and Doudna, J.A. (1996) RNA tertiary structure mediation by adenosine platforms. Science 273, 1696–1699.

[38] Golden, B.L., Gooding, A.R., Podell, E.R. and Cech, T.R. (1998) A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme. Science 282, 259–264.

[39] Roman, J. and Woodson, S.A. (1998) Integration of the *Tetrahymena* group I intron into bacterial rRNA by reverse splicing in vivo. Proc. Natl. Acad. Sci. USA 95, 2134–2139.

[40] Puttaraju, M. and Been, M.D. (1992) Group I permuted intron-exon (PIE) sequences self-splice to produce circular exons. Nucleic Acids Res. 20, 5357–5364.

[41] Puttaraju, M., Perrotta, A.T. and Been, M.D. (1993) A circular trans-acting hepatitis delta virus ribozyme. Nucleic Acids Res. 21, 4253–4258.

[42] Puttaraju, M. and Been, M.D. (1996) Circular ribozymes generated in *Escherichia coli* using group I self-splicing permuted intron-exon sequences. J. Biol. Chem. 271, 26081–26087.

[43] Sullenger, B.A. and Cech, T.R. (1994) Ribozyme-mediated repair of defective mRNA by targeted, trans-splicing. Nature 371, 619–622.

[44] Jones, J.T., Lee, S.W. and Sullenger, B.A. (1996) Tagging ribozyme reaction sites to follow trans-splicing in mammalian cells. Nature Med. 2, 643–648.

[45] Lan, N., Howrey, R.P., Lee, S.W., Smith, C.A. and Sullenger, B.A. (1998) Ribozyme-mediated repair of sickle beta-globin mRNAs in erythrocyte precursors. Science 280, 1593–1596.

[46] Michel, F. and Ferat, J.L. (1995) Structure and activities of group II introns. Annu. Rev. Biochem. 64, 435–461.

[47] Jacquier, A. (1996) Group II introns: elaborate ribozymes. Biochimie 78, 474–487.

[48] Jarrell, K.A., Peebles, C.L., Dietrich, R.C., Romiti, S.L. and Perlman, P.S. (1988) Group II intron self-splicing. Alternative reaction conditions yield novel products. J. Biol. Chem. 263, 3432–3439.

[49] Schmelzer, C., Schmidt, C. and Schweyen, R.J. (1982) Identification of splicing signals in introns of yeast mitochondrial split genes: mutational alterations in intron bI1 and secondary structures in related introns. Nucleic Acids Res. 10, 6797–6808.

[50] Koch, J.L., Boulanger, S.C., Dib-Hajj, S.D., Hebbar, S.K. and Perlman, P.S. (1992) Group II introns deleted for multiple substructures retain self-splicing activity. Mol. Cell. Biol. 12, 1950–1958.

[51] Qin, P.Z. and Pyle, A.M. (1998) The architectural organization and mechanistic function of group II intron structural elements. Curr. Opin. Struct. Biol. 8, 301–308.

[52] Copertino, D.W. and Hallick, R.B. (1993) Group-II and group-III introns of twintrons – potential relationships with nuclear pre-messenger RNA introns. Trends Biochem. Sci. 18, 467–471.

[53] Bonen, L. (1993) Trans-splicing of pre-mRNA in plants, animals, and protists. FASEB J. 7, 40–46.

[54] Podar, M., Chu, V.T., Pyle, A.M. and Perlman, P.S. (1998) Group II intron splicing in vivo by first-step hydrolysis. Nature 391, 915–918.

[55] Morl, M., Niemer, I. and Schmelzer, C. (1992) New reactions catalyzed by a group II intron ribozyme with RNA and DNA substrates. Cell 70, 803–810.

[56] Mikheeva, S., Hakim-Zargar, M., Carlson, D. and Jarrell, K. (1997) Use of an engineered ribozyme to produce a circular human exon. Nucleic Acids Res. 25, 5085–5094.

[57] Darr, S.C., Brown, J.W. and Pace, N.R. (1992) The varieties of ribonuclease P. Trends Biochem. Sci. 17, 178–182.

[58] Altman, S., Kirsebom, L. and Talbot, S. (1993) Recent studies of ribonuclease P. FASEB J. 7, 7–14.

[59] Kirsebom, L.A. (1995) RNase P – A Scarlet Pimpernel. Mol. Microbiol. 17, 411–420.

[60] Pace, N.R. and Brown, J.W. (1995) Evolutionary perspective on the structure and function of ribonuclease P, a ribozyme. J. Bacteriol. 177, 1919–1928.

[61] Peck-Miller, K.A. and Altman, S. (1991) Kinetics of the pro-cessing of the precursor to 4.5S RNA, a naturally occurring substrate for RNase P from *Escherichia coli*. J. Mol. Biol. 221, 1–5.

[62] Karwan, R. (1993) RNase MRP/RNase P: a structure-function relation conserved in evolution? FEBS Lett. 319, 1–4.

[63] Morrissey, J.P. and Tollervey, D. (1995) Birth of the snoRNPs: the evolution of RNase MRP and the eukaryotic pre-rRNA-processing system. Trends Biochem. Sci. 20, 78–82.

[64] Chamberlain, J.R., Lee, Y., Lane, W.S. and Engelke, D.R. (1998) Purification and characterization of the nuclear RNase P holoenzyme complex reveals extensive subunit overlap with RNase MRP. Genes Dev. 12, 1678–1690.

[65] James, B.D., Olsen, G.J., Liu, J.S. and Pace, N.R. (1988) The secondary structure of ribonuclease P RNA, the catalytic element of a ribonucleoprotein enzyme. Cell 52, 19–26.

[66] Brown, J.W. and Haas, E.S. (1996) Ribonuclease P structure and function in Archaea. Mol. Biol. Rep. 22, 131–134.

[67] Chen, J.L. and Pace, N.R. (1997) Identification of the universally conserved core of ribonuclease P RNA. RNA 3, 557–560.

[68] Westhof, E. and Altman, S. (1994) 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.

[69] Harris, M.E., Nolan, J.M., Malhotra, A., Brown, J.W., Harvey, S.C. and Pace, N.R. (1994) Use of photoaffinity cross-linking and molecular modeling to analyze the global architecture of ribonuclease P RNA. EMBO J. 13, 3953–3963.

[70] Chen, J.L., Nolan, J.M., Harris, M.E. and Pace, N.R. (1998) Comparative photocross-linking analysis of the tertiary structures of *Escherichia coli* and *Bacillus subtilis* RNase P RNAs. EMBO J. 17, 1515–1525.

[71] Massire, C., Jaeger, L. and Westhof, E. (1998) Derivation of the three-dimensional architecture of bacterial ribonuclease P RNAs from comparative sequence analysis. J. Mol. Biol. 279, 773–793.

[72] Altman, S. (1995) RNase P in research and therapy. BioTechnology 13, 327–329.

[73] Liu, F. and Altman, S. (1995) Inhibition of viral gene expression by the catalytic RNA subunit of RNase P from *Escherichia coli*. Genes Dev. 9, 471–480.

[74] Symons, R.H. (1992) Small catalytic RNAs. Annu. Rev. Biochem. 61, 641–671.

[75] Chowrira, B.M., Pavco, P.A. and McSwiggen, J.A. (1994) In vitro and in vivo comparison of hammerhead, hairpin, and hepatitis delta virus self-processing ribozyme cassettes. J. Biol. Chem. 269, 25856–25864.

[76] McKay, D.B. (1996) Structure and function of the hammerhead ribozyme: an unfinished story. RNA 2, 395–403.

[77] Stage-Zimmermann, T.K. and Uhlenbeck, O.C. (1998) Hammerhead ribozyme kinetics. RNA 4, 875–889.

[78] Wedekind, J.E. and McKay, D.B. (1998) Crystallographic structures of the hammerhead ribozyme: relationship to ribozyme folding and catalysis. Annu. Rev. Biophys. Biomol. Struct. 27, 475–502.

[79] Doudna, J.A. and Cate, J.H. (1997) RNA structure: crystal clear? Curr. Opin. Struct. Biol. 7, 310–316.

[80] Murray, J.B., Terwey, D.P., Maloney, L., Karpeisky, A., Usman, N., Beigelman, L. and Scott, W.G. (1998) The structural basis of hammerhead ribozyme self-cleavage. Cell 92, 665–673.

[81] Scott, W.G., Murray, J.B., Arnold, J.R.P., Stoddard, B.L. and Klug, A. (1996) Capturing the structure of a catalytic RNA intermediate: the hammerhead ribozyme. Science 274, 2065–2069.

[82] Burke, J.M. (1996) Hairpin ribozyme: current status and future prospects. Biochem. Soc. Trans. 24, 608–615.

[83] Earnshaw, D.J. and Gait, M.J. (1997) Progress toward the structure and therapeutic use of the hairpin ribozyme. Antisense Nucleic Acid Drug Dev. 7, 403–411.

[84] Hampel, A. (1998) The hairpin ribozyme: discovery, two-dimensional model, and development for gene therapy. Prog. Nucleic Acid Res. Mol. Biol. 58, 1–39.

[85] Earnshaw, D.J., Masquida, B., Muller, S., Sigurdsson, S.T., Eckstein, F., Westhof, E. and Gait, M.J. (1997) Inter-domain cross-linking and molecular modelling of the hairpin ribozyme. J. Mol. Biol. 274, 197–212.

[86] Walter, N.G., Hampel, K.J., Brown, K.M. and Burke, J.M. (1998) Tertiary structure formation in the hairpin ribozyme monitored by fluorescence resonance energy transfer. EMBO J. 17, 2378–2391.

[87] Moosbauer, J. and Tabler, M. (1997) A helix 1-extended hairpin ribozyme exhibits altered cleavage behavior in vitro. Antisense Nucleic Acid Drug Dev. 7, 79–87.

[88] Sargueil, B., Pecchia, D.B. and Burke, J.M. (1995) An improved version of the hairpin ribozyme functions as a ribonucleoprotein complex. Biochemistry 34, 7739–7748.

[89] Kennell, J.C., Saville, B.J., Mohr, S., Kuiper, M.T., Sabourin, J.R., Collins, R.A. and Lambowitz, A.M. (1995) The VS catalytic RNA replicates by reverse transcription as a satellite of a retroplasmid. Genes Dev. 9, 294–303.

[90] Beattie, T.L. and Collins, R.A. (1997) Identification of functional domains in the self-cleaving *Neurospora* VS ribozyme using damage selection. J. Mol. Biol. 267, 830–840.

[91] Rastogi, T. and Collins, R.A. (1998) Smaller, faster ribozymes reveal the catalytic core of *Neurospora* VS RNA. J. Mol. Biol. 277, 215–224.

[92] Guo, H.C. and Collins, R.A. (1995) Efficient trans-cleavage of a stem-loop RNA substrate by a ribozyme derived from neurospora VS RNA. EMBO J. 14, 368–376.

[93] Olive, J.E. and Collins, R.A. (1998) Spermine switches a *Neu-rospora* VS ribozyme from slow Cis cleavage to fast trans cleavage. Biochemistry 37, 6476–6484.

[94] Taylor, J.M. (1992) The structure and replication of hepatitis delta virus. Annu. Rev. Microbiol. 46, 253–276.

[95] Lai, M.M.C. (1995) The molecular biology of hepatitis delta virus. Annu. Rev. Biochem. 64, 259–286.

[96] Lazinski, D.W. and Taylor, J.M. (1995) Regulation of the hepatitis delta virus ribozymes to cleave or not to cleave? RNA 1, 225–233.

[97] Tanner, N.K. (1995) The catalytic RNAs from hepatitis delta virus: structure, function and applications. In: The Unique Hepatitis Delta Virus (Dinter-Gottlieb, G., Ed.), pp. 11–31. R.G. Landes, Austin, TX.

[98] Been, M.D. and Wickham, G.S. (1997) Self-cleaving ribozymes of hepatitis delta virus RNA. Eur. J. Biochem. 247, 741–753.

[99] Tanner, N.K. (1998) Biochemistry of the hepatitis delta virus catalytic RNAs. In: Ribozymes in Gene Therapy of Cancer (Scanlon, K.J. and Kashani-Sabet, M., Eds.). R.G. Landes, Austin, TX.

[100] Tanner, N.K., Schaff, S., Thill, G., Petit-Koskas, E., Crain-Denoyelle, A.M. and Westhof, E. (1994) A three-dimensional model of hepatitis delta virus ribozyme based on biochemical and mutational analyses. Curr. Biol. 4, 488–498.

[101] Bravo, C., Lescure, F., Laugaa, P., Fourrey, J.L. and Favre, A. (1996) Folding of the HDV antigenomic ribozyme pseudoknot structure deduced from long-range photocrosslinks. Nucleic Acids Res. 24, 1351–1359.

[102] Branch, A.D. and Polaskova, J.A. (1995) 3-D models of the antigenomic ribozyme of the hepatitis delta agent with eight new contacts suggested by sequence analysis of 188 cDNA clones. Nucleic Acids Res. 23, 4180–4189.

[103] Ferre-D'Amare, A.R., Zhou, K. and Doudna, J.A. (1998) Crystal structure of a hepatitis delta virus ribozyme. Nature 395, 567–574.

[104] Hertel, K.J., Pardi, A., Uhlenbeck, O.C., Koizumi, M., Ohtsuka, E., Uesugi, S., Cedergren, R., Eckstein, F., Gerlach, W.L., Hodgson, R. and Symons, R.H. (1992) Numbering system for the hammerhead. Nucleic Acids Res. 20, 3252.

[105] Siwkowski, A., Shippy, R. and Hampel, A. (1997) Analysis of hairpin ribozyme base mutations in loops 2 and 4 and their effects on cis-cleavage in vitro. Biochemistry 36, 3930–3940.