Annu. Rev. Biophys. Biomol. Struct. 2001. 30:457–75
Reprinted from Annu. Rev. Biochem. 2000. 69:597–615

RIBOZYME STRUCTURES AND MECHANISMS

Elizabeth A. Doherty
Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut, 06520; e-mail: doherty@csb.yale.edu

Jennifer A. Doudna
Department of Molecular Biophysics and Biochemistry, and Howard Hughes Medical Institute, Yale University, New Haven, Connecticut, 06520; e-mail: doudna@csb.yale.edu

Key Words ribozyme, RNA structure, RNA catalysis, RNA self-cleavage, RNA self-splicing

■ **Abstract** The past few years have seen exciting advances in understanding the structure and function of catalytic RNA. Crystal structures of several ribozymes have provided detailed insight into the folds of RNA molecules. Models of other biologically important RNAs have been constructed based on structural, phylogenetic, and biochemical data. However, many questions regarding the catalytic mechanisms of ribozymes remain. This review compares the structures and possible catalytic mechanisms of four small self-cleaving RNAs: the hammerhead, hairpin, hepatitis delta virus, and in vitro–selected lead-dependent ribozymes. The organization of these small catalysts is contrasted to that of larger ribozymes, such as the group I intron.

CONTENTS

INTRODUCTION 458
RIBOZYME CATALYSIS 459
SMALL SELF-CLEAVING RNAs 460
Hammerhead Ribozyme Structure 460
Hammerhead Ribozyme Catalysis 461
Leadzyme Motif Structure and Catalysis 463
Hepatitis Delta Virus Ribozyme Structure 463
Hepatitis Delta Virus Ribozyme Catalysis 465
Hairpin Ribozyme Structural Models and Catalysis 465
Organization of Small Self-Cleaving RNAs 467
ARCHITECTURE OF A COMPLEX RIBOZYME: The *Tetrahymena thermophila* Group I Ribozyme 467
Structural Models of the *Tetrahymena* Ribozyme 468
Peripheral Elements and Maintenance of the Active Structure 470
COMPARATIVE ORGANIZATION OF CATALYTIC RNAs 471

INTRODUCTION

Seven naturally occurring classes of catalytic RNA have been identified to date, all of which catalyze cleavage or ligation of the RNA backbone by transesterification or hydrolysis of phosphate groups. The hammerhead, hepatitis delta virus (HDV), hairpin, and *Neurospora* Varkud satellite (VS) ribozymes are small RNAs of 50–150 nucleotides that perform site-specific self-cleavage (1–7). Found in viral, virusoid, or satellite RNA genomes, they process the products of rolling circle replication into genome-length strands (8). The general mechanism of these ribozymes is similar to that of many protein ribonucleases in which a 2′ oxygen nucleophile attacks the adjacent phosphate in the RNA backbone, resulting in cleavage products with 2′,3′-cyclic phosphate and 5′ hydroxyl termini (Figure 1). Unlike protein ribonucleases, however, ribozymes cleave only at a specific location, using base-pairing and tertiary interactions to help align the cleavage site within the catalytic core. The evolutionary maintenance of these sequences may be due to their site specificity and to the simplicity and efficiency of genome self-cleavage.

Group I and II introns and ribonuclease P (RNase P) are larger, more structurally complex ribozymes several hundred nucleotides in length (9–11). RNase P cleaves precursor RNA substrates at specific sites to generate functional 5′ termini (12), and group I and II introns catalyze two-step self-splicing reactions (13–15). In these large ribozymes, the nucleophile and the labile phosphate are located on different molecules or are greatly separated in sequence. Thus, the complex folds of these RNAs serve to orient the nucleophile and phosphate to ensure accurate cleavage or splicing.

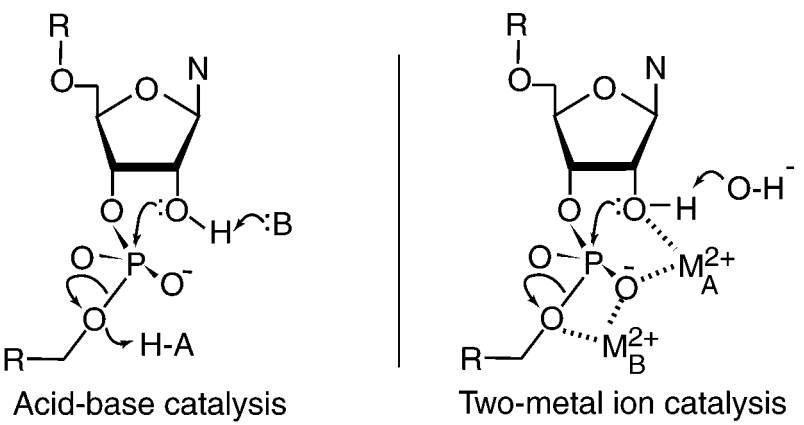

Figure 1 RNA cleavage by acid-base catalysis and two-metal ion catalysis. :B is a general base and H-A is a general acid. *Thick dashed lines* show direct inner-sphere coordination between divalent cations M_A and M_B and oxygen atoms. N represents a purine or pyrimidine base and R denotes chain continuation.

In the last several years, crystal structures of the hammerhead, hepatitis delta virus, and group I ribozymes have been determined, providing detailed views of the tertiary folds of these RNAs. The crystal structure of an in vitro–selected self-cleaving ribozyme called the leadzyme has also been determined. Models of the hairpin ribozyme have been constructed based on nuclear magnetic resonance (NMR) structures of separated domains and biochemical data. Furthermore, recent results suggest that the mechanisms used by RNA catalysts are more varied than previously thought. In this review we compare and contrast the overall organization of these ribozymes and discuss their catalytic mechanisms. This review does not cover RNA tertiary structural motifs, which have been extensively reviewed elsewhere (16–19).

RIBOZYME CATALYSIS

The notion that RNA could be catalytic was initially surprising since it lacks the diversity of functional groups characteristic of protein enzymes. Additionally, whereas protein secondary structure puts amino acid side chains on the outside of α-helices and β-sheets, optimally located for tertiary contacts, in RNA the unique chemical groups of the bases are largely on the interior of base-paired duplexes. Thus, RNA packing into stable tertiary structures was thought to be difficult because of the limited number of bases available for tertiary contacts, and the high density of negative charges and flexibility of the phosphate backbone. Despite these potential challenges, RNA enzymes are capable of accelerating phosphoryl transfer reactions by 10^5- to 10^{11}-fold over background rates (20–23). In addition, ribozymes selected in vitro can accelerate a variety of reactions related to phosphate transesterification, including acyl transfer (24–29).

How do ribozymes accomplish significant rate enhancements? RNA hydrolysis by protein ribonucleases such as RNase A occurs via acid-base catalysis in which one histidine imidazole group, acting as a general base, abstracts a proton from the 2′ hydroxyl nucleophile while a second histidine, acting as a general acid, donates a proton to the 5′ hydroxyl leaving group (Figure 1*a*). The penta-coordinated phosphate of the transition state is thought to be stabilized by the positive charge of a nearby lysine side chain (30). It is the precise positioning of these groups and the ability of the protein to modulate the pK_as of the reactive histidines that are responsible for the rate acceleration by RNase A. In contrast, RNA lacks functional groups with pK_as near physiological pH, potentially precluding similar mechanisms of acid-base catalysis (31). Since Mg^{2+} or other divalent cations are essential for the folding of many RNAs (32, 33), and RNA structures often contain specific binding sites for divalent ions (for examples see 34–36), ribozymes may carry out metal-ion–assisted catalysis that is similar to the action of several protein enzymes which catalyze phosphate chemistry (Figure 1*b*) (33, 37–39).

Divalent metal ions, particularly Mg^{2+}, could play several roles in catalysis. A metal ion coordinated to a hydroxide might activate a hydroxyl or water nucleophile by deprotonation, or a divalent ion might directly coordinate the nucleophilic

oxygen, making the oxygen more susceptible to deprotonation by hydroxide ions. Metal ions also might stabilize the transition state by direct inner-sphere coordination to the pentavalent scissile phosphate group and might stabilize the leaving group by protonating or directly coordinating the leaving oxygen atom. Finally, metal ions might also stabilize the transition state structure by donating positive charge.

Alkaline phosphatase and DNA polymerase I use two divalent cations positioned about 4 Å apart in catalysis, the first to activate the nucleophile and the second to stabilize the transition state and leaving group (37, 38). It has been proposed that a similar two-metal ion mechanism occurs in RNA catalysis (Figure 1*b*) (33, 39). This proposal has been tested in several systems by switching the metal ion specificity of phosphate groups in the RNA backbone, particularly at the labile phosphates. Sulfur substitution of phosphate oxygens reduces catalytic activity when the corresponding oxygen forms a critical hydrogen bond or binds a magnesium ion by direct inner-sphere coordination (40). Rescue of less active mutants by addition of a thiophilic metal ion such as Mn^{2+} indicates that the corresponding oxygen atom directly coordinates a structurally or catalytically important Mg^{2+} ion (40). These experiments have allowed detection of magnesium binding sites in the catalytic cores of several ribozymes (41–57). However, direct coordination of metal ions to nucleophilic, transition state or leaving group oxygens has not been inferred in all systems, suggesting that some ribozymes may catalyze phosphate chemistry through different mechanisms, as discussed below.

SMALL SELF-CLEAVING RNA*s*

The first X-ray crystal structure of a catalytic RNA was that of the self-cleaving hammerhead ribozyme (58, 59). Recently the crystal structure of the genomic strand HDV ribozyme (60) and solution and crystal structures of the in vitro–selected leadzyme motif (61, 62) have been determined. Furthermore, structures of the isolated domains of the hairpin ribozyme have been determined by solution NMR (63, 64), and models of the hairpin active site have been developed using a combination of NMR and biochemical data (65, 66). Below we review the structure and catalytic mechanism of each of these ribozymes.

Hammerhead Ribozyme Structure

The hammerhead ribozyme was initially discovered as a self-cleaving sequence within small RNA satellites of plant viruses (4, 6, 7). The minimal functional RNA consists of three short helices and a universally conserved junction sequence and can be made from two or more separate RNAs for cleavage in *trans* (Figure 2) (5, 67, 68).

Two crystal structures of noncleavable variants of the hammerhead ribozyme have been solved, one with a DNA substrate strand and the other with an RNA substrate containing a 2′ O-methyl group at the cleavage site (58, 59). Both show

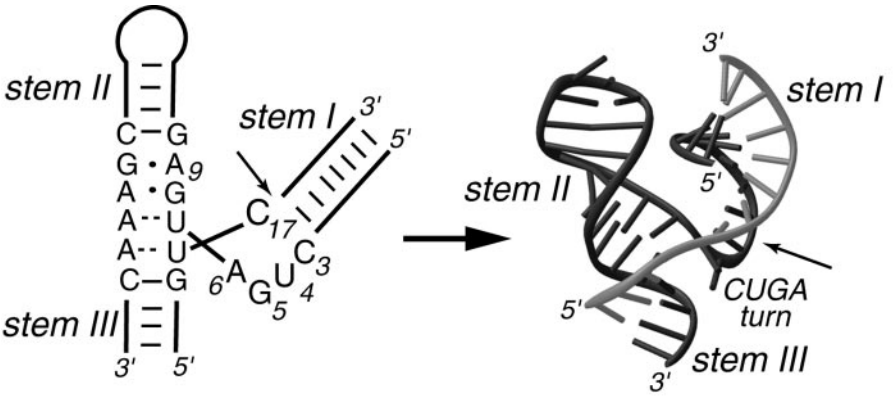

Figure 2 Secondary and tertiary structure of a hammerhead ribozyme. Universally conserved bases are shown as letters and those referred to in the text are numbered. *Dashed lines and dots* show non–Watson-Crick interactions between bases; *solid lines* represent Watson-Crick base pairs. The site of cleavage is marked with an *arrow*. (Adapted from Reference 58.)

the three helices arranged in a Y shape, as predicted by fluorescence and native gel electrophoresis data (Figure 2) (69, 70). Stems II and III are essentially coaxial, while helix I lies at a sharp angle to helix II. Backbone distortions at the junction of helices II and III force nucleotide C17 to stack on stem I rather than on stem III, placing it in the active site pocket at the three-helix junction. The scissile phosphate on the 3′ side of C17 lies above a hairpin turn of the backbone formed by the C3-A6 sequence. This CUGA turn is strikingly similar to that found in the anticodon loop of $tRNA^{Phe}$, which serves as a metal binding pocket (71). Although both structures revealed unforeseen details of the RNA fold, the mechanism of hammerhead catalysis remains unclear, as discussed below.

Hammerhead Ribozyme Catalysis

Many experiments have been performed to determine the mechanism of hammerhead self-cleavage and the role of divalent cations in the reaction. Sulfur substitution for the scissile phosphate oxygens showed that there is inversion of configuration about the phosphate during the reaction, indicating that the reaction proceeds by in-line attack of the nucleophile (42, 72, 73). Further, in most of the experiments in which sulfur was substituted for the pro-Rp[1] nonbridging oxygen

[1]Phosphate groups in the RNA backbone are pro-chiral, and the two nonbridging oxygen atoms of these groups are termed the pro-Rp and the pro-Sp oxygens. The pro-Rp oxygen is easily substituted with sulfur by in vitro transcription of RNA in the presence of sulfur containing nucleoside triphosphates. Substitution of the pro-Sp oxygen with sulfur requires chemical synthesis and is thus rarely performed.

of the scissile phosphate, the reduced activity was rescued by addition of a thiophilic metal ion, suggesting that a metal ion directly coordinates this oxygen in the transition state (41–43, 45, 46, 73). The reaction rate increases linearly with pH, indicating that the nucleophile is activated by a hydroxide ion (74). Either a metal ion hydroxide deprotonates the 2′ hydroxyl directly, or a metal ion coordinated to the 2′ hydroxyl increases the acidity of the 2′ oxygen, rendering it susceptible to attack by a hydroxide ion from solution (74, 75). The ion coordinated to the pro-Rp phosphate oxygen could perform either of these functions (Figure 1*b*). If the hammerhead uses a two-metal ion mechanism for self-cleavage, then an additional directly coordinated ion should stabilize the leaving group oxygen (Figure 1*b*) (33, 39). While sulfur substitution for the leaving group oxygen has failed to identify such a metal ion, other observations support its existence (44, 75, 76). Thus, there is debate over the number of divalent cations directly involved in hammerhead catalysis. Strikingly, recent experiments demonstrate that the ribozyme can function in the complete absence of divalent cations at extremely high ionic strengths (77). This result shows that divalent cations are not essential cofactors in the reaction, though at least one directly coordinated magnesium ion appears to be involved in catalysis under most conditions in vitro.

Unfortunately, determination of the hammerhead crystal structure did little to settle the debate over the role of metal ions in catalysis. While several divalent metal ions were identified by soaking Mn^{2+} into hammerhead ribozyme crystals (58, 59), none of these ions was close enough to the scissile phosphate to play a direct role in catalysis without requiring a precleavage conformational change. Therefore, the crystal structure of an unmodified hammerhead was determined by cryo-cooling the crystals to trap the ribozyme just prior to catalysis. This structure revealed a metal ion directly coordinated to the pro-Rp nonbridging oxygen of the scissile phosphate, as expected from biochemical data (78). However, no metal ion was found near the leaving group in any subsequent structure.

Several observations strongly suggest that the original crystal structures represented an initial ground state of the ribozyme that must undergo a conformational change before catalysis. A number of catalytically important base and backbone functional groups identified by mutagenesis make only limited contacts in the crystal structures (79). Furthermore, one of the metal binding sites located near the three-helix junction was subsequently found to inhibit catalysis when occupied (80). Finally, previous studies demonstrated that the reaction proceeds by in-line attack of the 2′ oxygen nucleophile, but the labile phosphate in both the original and freeze-trapped crystal structures was not correctly oriented for nucleophilic attack (42, 72, 73, 78). To visualize the hammerhead ribozyme in a reactive conformation, a mutation was introduced adjacent to the leaving oxygen atom to kinetically trap the ribozyme immediately prior to bond breaking (81). In the crystal structure of this modified RNA, a rotation of the scissile phosphate and base C17 was observed, consistent with an in-line attack mechanism. There was also a change in the conformation of the A9 phosphate, consistent with the observation that a metal

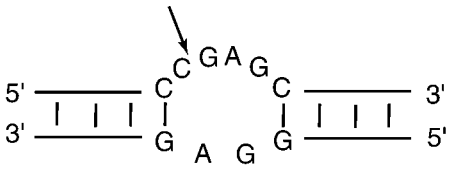

Figure 3 Secondary structure of the in vitro–selected leadzyme motif. Two Watson-Crick helices flank an internal loop containing the cleavage site (at *arrow*). Conserved bases are shown as *letters*.

ion coordinated to this phosphate binds with higher affinity in the transition state than in the ground state (82). Strikingly, the architecture of the three-helix junction was not grossly perturbed by the backbone rotation at the cleavage site. This result demonstrates the high degree of flexibility within the hammerhead active site.

Leadzyme Motif Structure and Catalysis

The observation that a lead-hydroxide ion cleaves tRNA at a specific site led to the use of in vitro selection to define the leadzyme, a small, lead-dependent self-cleaving RNA (83–85). The leadzyme consists of two short Watson-Crick duplexes flanked by an internal loop that accelerates site-specific RNA cleavage by ∼1000-fold (Figure 3) (84, 85). The reaction is highly specific for Pb^{2+} ions and proceeds by a cleavage mechanism similar to that of other self-cleaving RNAs (86–88). The lead-hydroxide ion is proposed to activate the 2′ hydroxyl nucleophile by deprotonation (86). Both NMR and crystal structures of this motif have recently been solved (61, 62). While the conformation of the internal loop varies, the labile phosphate is not positioned for in-line attack by the nucleophile in either structure. NMR studies also demonstrated that the internal loop is dynamic (89). Thus, like the hammerhead ribozyme active site, the leadzyme active site is conformationally flexible.

Hepatitis Delta Virus Ribozyme Structure

The biological function of the HDV ribozyme, like that of the hammerhead ribozyme, is to cleave rolling circle replication products into genome-length units (2, 3). HDV replication produces multimers of antigenomic and genomic RNAs, both of which contain similar ribozymes. The HDV ribozyme is the fastest naturally occurring RNA catalyst, capable of self-cleavage at a rate ∼100-fold faster than that of the hammerhead (90, 91). This ribozyme has a number of unusual properties compared with other self-cleaving RNAs. It is extremely stable, with an optimal reaction temperature of about 65°C, and is also reactive at temperatures as high as 80°C and in buffers containing 5 M urea or 18 M formamide (90–92). Furthermore, the substrate strand upstream of the cleavage site is not base-paired to the ribozyme. Thus, cleavage depends only on the presence of a single nucleotide 5′ of the labile phosphate. Finally, though the presence of divalent cations is

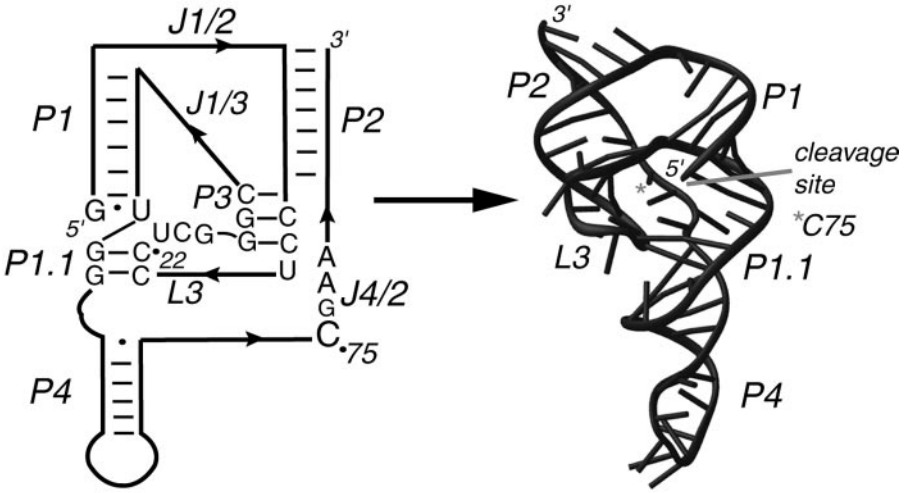

Figure 4 Secondary and tertiary structure of the hepatitis delta virus genomic cleavage product. Important bases are shown as *letters* and those referred to in the text are *numbered. P1–P4* denote base-paired stems, *L* denotes loops, and *J* denotes joining regions between two helices. (Adapted from Reference 60.)

required for catalysis, self-cleavage is efficient with a variety of different ions and at much lower concentrations than those required for the activity of other catalytic RNAs (77, 92, 93).

The crystal structure of the genomic cleavage product revealed the basis for the extreme stability of this ribozyme (60). The HDV ribozyme is folded into a double pseudoknot containing five helical stems, P1–P4 (Figure 4). The P1 duplex is coaxially stacked upon P1.1 and P4, while P2 is stacked on P3. The two stacks are positioned side by side, linked by five strand crossovers and further constrained by the P1.1 pairing. Intriguingly, no well-ordered metal ions were detected either in the electron density maps from the original experiments or in maps calculated using data measured from crystals soaked in Mn^{2+} (60, 94). This observation suggests that specifically bound divalent cations are not required to stabilize the structure of the cleavage product. Unlike the hammerhead ribozyme and the leadzyme, in which the active site is solvent exposed, the 5′ hydroxyl leaving group of the HDV ribozyme is deeply buried within the fold, surrounded by functionally important residues from the P3 stem and the L3, J1/3, and J4/2 segments (Figure 4). The observations that only a single nucleotide 5′ of the scissile phosphate is required for catalysis in this ribozyme, and that functional groups determined to be important through cross-linking and mutagenesis studies surround the 5′ hydroxyl leaving group in the crystal structure, indicate that the product and transition-state structures are similar (56, 90, 95–98). In addition to biochemical data, the burial of the

active site and lack of metal ions in the crystal structure provide hints that the HDV cleavage mechanism may be distinct from that of the hammerhead ribozyme and leadzyme.

Hepatitis Delta Virus Ribozyme Catalysis

The observation that low concentrations of divalent cations are required for HDV ribozyme function indicates that they play an important role in folding or catalysis (77). Yet several lines of evidence suggest that their catalytic role is different from that in the hammerhead ribozyme. First, sulfur substitution reveals that only two nonbridging pro-Rp phosphate oxygens, at the scissile phosphate and at C22, are critical for catalysis. Both mutations are insensitive to Mn^{2+}, however, implying that these oxygens are not directly coordinated to metal ions but make other key interactions (56). In fact, the pro-Rp phosphate oxygen at C22 participates in a network of hydrogen bonds to the catalytically important base C75 in the crystal structure (60). Second, the pH dependence of the reaction suggests that an RNA functional group, rather than a hydroxide ion, activates the nucleophile (see below) (98–100). Third, the rate of self-cleavage is equivalent with Mg^{2+} and Ca^{2+}, and is also efficient in Mn^{2+} and even Sr^{2+}. This suggests that divalent cations bind to a site with flexible geometry and thus low affinity (56). The presence of a metal ion at the active site of the ribozyme is implied by studies of the cleavage of a 2′,5′-modified phosphate linkage (101). Binding sites for divalent ions were recently identified in the J4/2 section of the genomic HDV ribozyme (102) and the P2 stem of the antigenomic ribozyme (103), but there is no evidence that these ions play direct roles in catalysis.

The HDV ribozyme crystal structure reveals that a base might play a key role in the reaction mechanism. Cytosine 75 is located extremely close to the 5′ hydroxyl leaving group in the structure. This base is excluded from solvent in a region of low electrostatic potential, surrounded by several negatively charged phosphate groups. This could increase the pK_a of C75, normally ~4 for the N3 imino group. Thus it has been proposed that C75 acts as a general base to activate the 2′ oxygen nucleophile, or as a general acid to protonate the leaving group oxygen (see Figure 1*a*) (60, 94). In support of this hypothesis, mutation of C75 to any other base is the most severe HDV ribozyme mutant, completely abolishing catalysis (96). In the case of a C to U mutation at the corresponding position in the antigenomic ribozyme, functionality is rescued by addition of imidazole, which may function as a general acid or base (100). Furthermore, the pH dependence of the reaction indicates that an RNA functional group with pK_a of ~6 is important for catalysis (100).

Hairpin Ribozyme Structural Models and Catalysis

The hairpin ribozyme, like the hammerhead ribozyme, is found within RNA satellites of plant viruses, where it performs a reversible self-cleavage reaction to

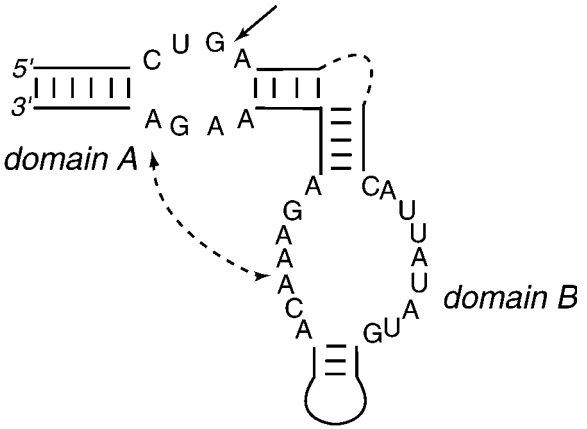

Figure 5 Secondary structure of a minimal hairpin ribozyme construct. Conserved nucleotides within domains A and B are *lettered*. The cleavage site is at the *arrow*. The *dashed line* at the junction of the two domains can vary in sequence and length. Tertiary folding causes the two helical domains to pack side by side, stabilized by tertiary interactions between the conserved nucleotides in the internal loops. The cleavage site is proposed to pack against portions of the domain B internal loop, rendering it inaccessible to solvent. (Adapted from Reference 66.)

process the products of rolling circle genome replication (7, 104–106). Models of the catalytic core of the hairpin ribozyme have been developed based on extensive mutagenesis, chemical probing, and NMR structures of the isolated helical domains (63–66). Chemical footprinting, cross-linking, and fluorescence experiments show that the two domains of the hairpin catalytic core interact side by side to create a buried active site, as seen in the HDV ribozyme (Figure 5) (65, 107, 108). However, in the hairpin ribozyme the helix-helix interaction is mediated by tertiary contacts between the domains rather than by covalent strand crossovers, and thus is likely to be more flexible.

Several lines of evidence suggest that divalent cations may not play a central role in catalysis for the hairpin ribozyme. Changes in pH and substitution of sulfur for either of the nonbridging oxygens of the scissile phosphate have minimal effects on the rate of the reaction under a wide variety of conditions (53, 54, 109). The reaction also proceeds efficiently in cobalt hexammine, an ion similar in size to a fully hydrated magnesium ion but which cannot directly coordinate RNA functional groups. All of these results strongly suggest that neither a metal hydroxide nor a directly coordinated metal ion is required for activity under most conditions (53, 54, 109). Furthermore, the hairpin ribozyme can also function when supplied with monovalent cations alone (77). Perhaps in the hairpin ribozyme, as is apparently the case in the HDV ribozyme, nucleotide bases play critical catalytic roles.

Organization of Small Self-Cleaving RNAs

Examination of the ribozyme structures discussed above reveals two general trends in the organization of self-cleaving RNAs. The hammerhead ribozyme and leadzyme active sites are exposed to solvent and coated with divalent cations, some of which may play direct roles in catalysis. The RNAs are flexible, allowing for conformational changes during the reaction. Since the solution and crystal structures of these RNAs are not in active conformations, it is likely that the most active conformers of these ribozymes are not the most thermodynamically stable and thus are present only transiently in solution. Although conformationally dynamic, the hammerhead ribozyme is an effective catalyst in part because the three-helix junction of the cleavage product is more flexible than that of the ground state, allowing entropy to drive the reaction forward by preventing religation of the product strands (21).

In contrast, the HDV ribozyme active site is buried from solvent within a compact fold stabilized by covalent attachment of the two helical stacks. The structural rigidity of the ribozyme might restrict the orientation of the labile phosphate, thereby allowing for a faster rate of catalysis. The reaction rate may be further augmented by the direct involvement of an active site cytosine in the catalytic mechanism. The hairpin ribozyme shares several organizational features with the HDV ribozyme. The active site is buried within a set of helices stacked side by side (65, 107, 108). However, the hairpin ribozyme fold is probably more dynamic than that of the HDV ribozyme since it is stabilized by tertiary contacts rather than by covalent attachments. The catalytic mechanism does not appear to involve directly coordinated divalent metal ions and could perhaps directly involve one or more bases.

ARCHITECTURE OF A COMPLEX RIBOZYME: The *Tetrahymena thermophila* Group I Ribozyme

Group I introns are considerably larger and more structurally complex than any of the self-cleaving RNAs. They are found in precursor mRNA, tRNA, and rRNA transcripts in a variety of organisms and self-splice in two steps to ligate flanking 5′ and 3′ exons, producing mature RNA transcripts (see Figure 7). Several hundred examples have now been identified, all of which share a common secondary structure and presumably a common reaction mechanism.

The *Tetrahymena thermophila* rRNA intron was the first group I intron discovered and is the best characterized (110). It is 421 nucleotides long, composed of a universally conserved catalytic core of roughly 200 nucleotides surrounded by several less conserved peripheral segments (Figure 6) (111). The ribozyme derived from the intron performs the first step of self-splicing with multiple turnover using an oligonucleotide to mimic the 5′ exon (Figures 6 and 7). The 3′ oxygen of an exogenous guanosine serves as the nucleophile. Experiments in which key

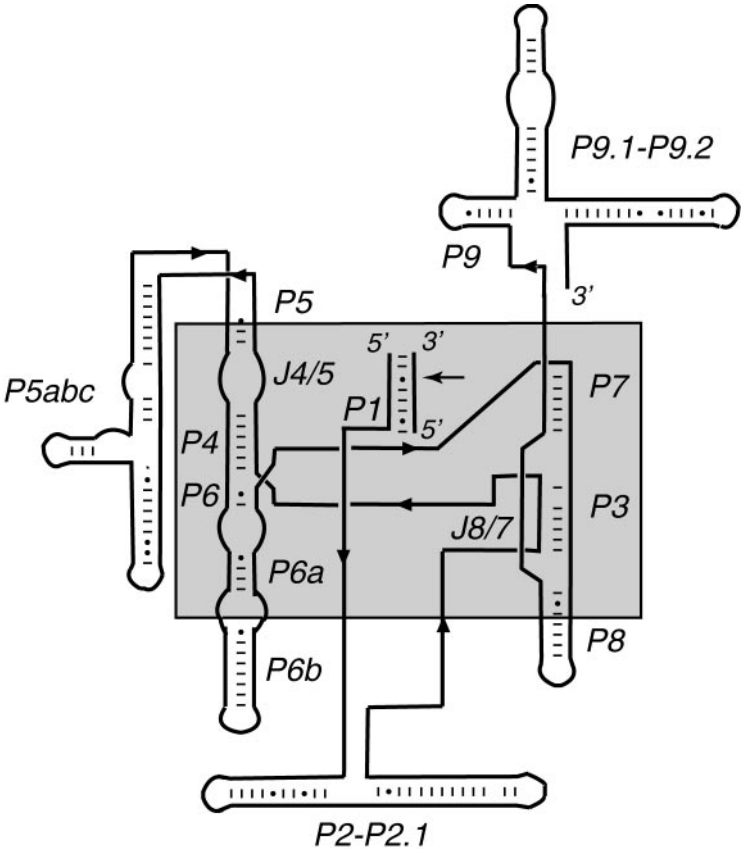

Figure 6 Schematic diagram of the *Tetrahymena thermophila* group I ribozyme secondary structure. The *P*, *L*, and *J* nomenclature is as described for Figure 4. The universally conserved catalytic core is inside the *gray box*, and the *P5abc*, *P2-P2.1*, and *P9.1-P9.2* peripheral elements are outside the box. *Lines with arrowheads* show connectivities between different secondary structural elements. The site of 5′ splicing is marked with an *arrow*.

oxygens were substituted by sulfur or nitrogen provided strong evidence that three directly coordinated Mg^{2+} ions participate in the catalytic mechanism to activate the nucleophile and stabilize the leaving group (Figure 7) (49–51, 112).

Structural Models of the *Tetrahymena* Ribozyme

A variety of experiments showed that the *Tetrahymena* ribozyme is folded into a compact, globular structure and revealed important tertiary interactions within the catalytic core. Footprinting with hydroxyl radicals, which cleave solvent-accessible riboses in the RNA backbone, demonstrated that large segments of

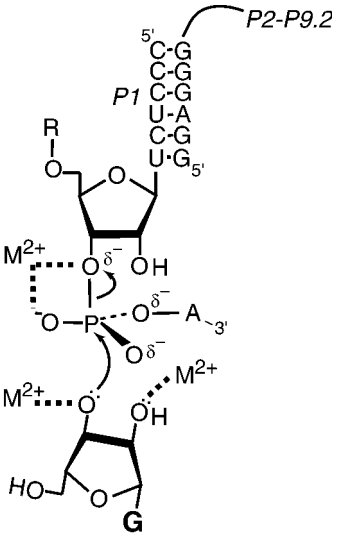

Figure 7 The *Tetrahymena* ribozyme reaction. The 3′ oxygen of an exogenous guanosine bound to the *P7* helix attacks the phosphate adjacent to a universally conserved base pair with G-U wobble in *P1*. Magnesium ions inferred from sulfur or nitrogen derivatization to be directly involved in catalysis are shown. (Adapted from References 50 and 112.)

the core are buried from solvent (32, 113). Phylogenetic comparison, combined with footprinting data, was used several years ago to construct a model of the arrangement of helices within the group I intron core (111). In this model the catalytic core is composed of two universally conserved helical stacks called P5-P4-P6 and P7-P3-P8, which form a cleft along which the P1 substrate helix binds (Figure 6). Mutagenesis and cross-linking experiments extended this model by independently locating the P1 and guanosine substrate binding sites and identifying tertiary interactions between the P5-P4-P6 and P7-P3-P8 segments (114–123). Further mutagenesis experiments and the 2.8-Å crystal structure of the P4-P6 domain (comprising the P5-P4-P6 stack and the P5abc peripheral region), revealed interactions between the core and peripheral elements leading to a model of the entire intron (120, 124).

Recently, a crystal structure of the *Tetrahymena* ribozyme core was solved at 5-Å resolution (125). The crystallization construct included the entire catalytic core of the ribozyme but lacked the P2-P2.1 and P9.1-P9.2 peripheral regions as well as the P1 substrate helix (Figure 6). The overall architecture of the ribozyme in the crystal structure is similar to that proposed in the phylogenetic models (111, 120). A model for the interaction of P1 and the catalytic core was developed from this structural data, in which P1 is in close proximity to the guanosine binding site on P7.

A detailed model of the *Tetrahymena* ribozyme active site has also been created using nucleotide analog interference mapping (Figure 8) (126). These assays identified tertiary contacts between the P1 helix and the flanking J4/5 and J8/7 regions (122, 127). The model constructed from these data also incorporates previously identified tertiary contacts between J8/7, part of a universally conserved pseudoknot (128) at the center of the catalytic core, and helices P4 (121), P1 (115),

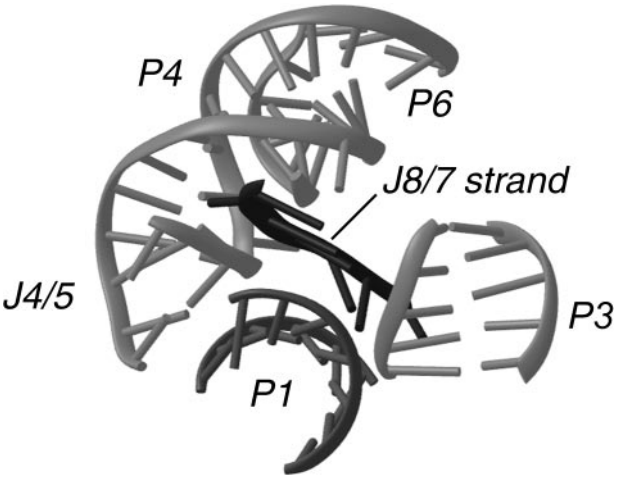

Figure 8 Active site of the *Tetrahymena* group I intron, showing the close positioning of several secondary structural elements. (Adapted from Reference 122.)

and P3 (111, 123). In the model, the J8/7 strand helps tightly pack the P5-P4-P6 and P7-P3-P8 helical stacks and forms an extensive triple-helix interaction with P1 (Figure 8) (122). Thus, P1 is tightly docked into the core in close proximity to several other elements of secondary structure.

Several of the tertiary contacts identified biochemically are not present in the 5-Å crystal structure. In particular, interactions between J8/7 and P1 are not compatible with the structure (115, 122), and a functionally critical base triple connecting J8/7 and P3 is absent, suggesting that the P3 helix in the structure must move 20–25 Å toward the P5- P4-P6 stack to achieve a conformation consistent with activity (123). Two possible explanations are that 5-Å resolution is insufficient to resolve nucleotide positions in the active site, or that the structure in the crystal represents a less active conformation. Thus, the P1 helix and P2-P2.1 and P9.1-P9.2 peripheral segments may be critical for proper organization of the active site, as discussed below.

Peripheral Elements and Maintenance of the Active Structure

Most group I introns contain at least one large structural element flanking the catalytic core, such as the P5abc, P2-P2.1, and P9.1-P9.2 segments in the *Tetrahymena* intron (Figure 6). These external appendages, termed peripheral elements, vary considerably in size, structure, and location, and can even be replaced by proteins (111, 120, 129). Despite this variability, the nearly universal presence of peripheral elements underscores their importance. Deletion of these segments generally reduces but does not eliminate self-splicing, suggesting that they are important but not essential for group I intron function (130–132).

The P5abc, P2-P2.1, and P9.1-P9.2 segments interact with each other to wrap around the outside of the *Tetrahymena* catalytic core (120). Results from chemical footprinting experiments show that deletion of one or more of these peripheral elements destabilizes the core, but deletion of segments of the core does not significantly affect the folding of the peripheral segments (133–135). Thus, these elements buttress a less stable catalytic core at the center of the globular fold.

The consequences of deletion of the P5abc peripheral element have been examined in detail (129–132, 136, 137). Data from footprinting experiments show that at nonphysiological concentrations of Mg^{2+}, removal of P5abc does not grossly perturb the global structure of the ribozyme. However, the catalytic core of a P5abc-deletion mutant ribozyme requires higher Mg^{2+} concentrations to fold and has subtle structural differences from the wild-type ribozyme core (136, 137). These differences are concentrated in the P7-P3 region near the substrate binding sites, on the opposite side of the catalytic core from the P5abc binding interface. Furthermore, the affinity of the guanosine and P1 substrates is severely compromised (137). Thus, the P5abc-deletion mutant ribozyme assembles into one or more less active structures. These results suggest that P5abc not only stabilizes the group I ribozyme core but also acts as an allosteric effector to organize the detailed architecture of the core from a distance, preventing the accumulation of misfolded structures.

COMPARATIVE ORGANIZATION OF CATALYTIC RNAs

The hammerhead, HDV, and group I ribozyme structures illustrate three different strategies to overcome the challenges of packing RNA into functional structures. The self-cleavage reaction is simple enough to allow an efficient catalyst to be created from a small RNA with a solvent-exposed active site, such as the hammerhead ribozyme. The three helices of this ribozyme do not pack closely except at the junction region, where they are correctly oriented by base-stacking and tertiary interactions involving a number of universally conserved bases. The structure of the leadzyme further demonstrates that a self-cleaving ribozyme does not require a true tertiary structure with closely packed RNA helices. However, the more rigid HDV ribozyme appears to enhance catalysis further by placing the active site between two helices packed side by side.

Group I introns catalyze a two-step reaction that requires positioning of two distinct substrates in the active site. Furthermore, a different set of substrates is used for each step of the reaction. The group I catalytic core must be stable enough to create high-affinity binding sites for each substrate, yet flexible enough to displace and reposition substrates between the two steps of the reaction. A rigid structure such as that of the HDV ribozyme would be incompatible with such a complex reaction. Instead, to stabilize and correctly position tertiary contacts between the elements of the catalytic core, group I introns contain a series of peripheral appendages that form a network of buttresses on the outside of the core.

Thus, group I introns consist of a loosely organized interior cinched together by a belt of exterior structural helices. Examination of the structure of group II introns and RNase P RNA suggests that these large ribozymes also contain core and peripheral domains similar to those in group I introns (138–140). The observation that group I peripheral elements can be replaced by proteins (129, 141) hints that an early role of proteins may have been to structurally support complex RNA enzymes such as a primitive ribosome. Indeed, recent structures of the ribosome and its components suggest that a major role of ribosomal proteins is to stabilize and orient critical sections of the RNA (142–147).

LITERATURE CITED

1. Saville B. 1990. *Cell* 61:685–96
2. Sharmeen L, Kuo MYP, Dinter-Gottlieb G, Taylor J. 1988. *J. Virol.* 62:2674–79
3. Kuo MYP, Sharmeen L, Dinter-Gottlieb G, Taylor J. 1988. *J. Virol.* 62:4439–44
4. Prody GA, Bakos JT, Buzayan JM, Schneider IR, Bruening G. 1986. *Science* 231:1577–80
5. Forster AC, Symons RH. 1987. *Cell* 49: 211–20
6. Hutchins CJ, Rathjen PD, Forster AC, Symons RH. 1986. *Nucleic Acids Res.* 14:3627–40
7. Buzayan JM, Gerlach WL, Bruening G. 1986. *Nature* 323:349–52
8. Branch AD, Robertson HD. 1984. *Science* 223:450–55
9. Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TR. 1982. *Cell* 31:147–57
10. Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S. 1983. *Cell* 35:849–57
11. Peebles C, Perlman P, Mecklenburg K, Petrillo M, Tabor J, et al. 1986. *Cell* 44: 213–33
12. Frank DN, Pace NR. 1998. *Annu. Rev. Biochem.* 67:153–80
13. Cech TR. 1990. *Annu. Rev. Biochem.* 59: 543–68
14. Michel F, Ferat J-L. 1995. *Annu. Rev. Biochem.* 64:435–61
15. Jacquier A. 1996. *Biochimie* 78:474–87
16. Westhof E, Masquida B, Jaeger L. 1996. *Fold. Des.* 1:R78–88
17. Moore PB. 1998. *Annu. Rev. Biophys. Biomol. Struct.* 27:35–58
18. Batey RT, Rambo RP, Doudna JA. 1999. *Angew. Chem. Int. Ed. Engl.* 38:2326–43
19. Ferre-D'Amare AR, Doudna JA. 1999. *Annu. Rev. Biophys. Biomol. Struct.* 28: 57–73
20. Herschlag D, Cech TR. 1990. *Biochemistry* 29:10159–71
21. Hertel KJ, Herschlag D, Uhlenbeck OC. 1994. *Biochemistry* 33:3374–85
22. Beebe JA, Fierke CA. 1994. *Biochemistry* 33:10294–304
23. Hegg LA, Fedor MJ. 1995. *Biochemistry* 34:15813–28
24. Wilson C, Szostak JW. 1995. *Nature* 374: 777–82
25. Illangasekare M, Sanchez G, Nickles T, Yarus M. 1995. *Science* 267:643–47
26. Lohse PA, Szostak JW. 1996. *Nature* 381: 442–44
27. Conn MM, Prudent JR, Schultz PG. 1996. *J. Am. Chem. Soc.* 118:7012–13
28. Zhang B, Cech TR. 1997. *Nature* 390: 96–100
29. Tarasow TM, Tarasow SL, Eaton BM. 1997. *Nature* 389:54–57
30. Richards FM, Wyckoff HW. 1971. In *Bovine Pancreatic Ribonuclease*, ed. PD Boyer, pp. 647–806. New York: Academic

31. Narlikar GJ, Herschlag D. 1997. *Annu. Rev. Biochem.* 66:19–59
32. Latham JA, Cech TR. 1989. *Science* 245: 276–82
33. Pyle AM. 1993. *Science* 261:709–14
34. Cate JH, Doudna JA. 1996. *Structure* 4: 1221–29
35. Cate JH, Hanna RL, Doudna JA. 1997. *Nat. Struct. Biol.* 4:553–58
36. Basu S, Rambo RP, Strauss-Soukup J, Cate JH, Ferre-D'Amare AR, et al. 1998. *Nat. Struct. Biol.* 5:986–92
37. Kim ES, Wyckoff HW. 1991. *J. Mol. Biol.* 218:449–64
38. Beese LS, Steitz TA. 1991. *EMBO J.* 10: 25–33
39. Steitz TA, Steitz JA. 1993. *Proc. Natl. Acad. Sci. USA* 90:6498–502
40. Eckstein F. 1979. *Acc. Chem. Res.* 12: 204–10
41. Ruffner DE, Uhlenbeck OC. 1990. *Nucleic Acids Res.* 18:6025–29
42. Koizumi M, Ohtsuka E. 1991. *Biochemistry* 30:5145–50
43. Dahm SC, Uhlenbeck OC. 1991. *Biochemistry* 30:9464–69
44. Kuimelis RG, McLaughlin LW. 1995. *J. Am. Chem. Soc.* 117:11019–20
45. Zhou D, Kumar PKR, Zhang L, Taira K. 1996. *J. Am. Chem. Soc.* 118:8969–70
46. Scott EC, Uhlenbeck OC. 1999. *Nucleic Acids Res.* 27:479–84
47. Ortoleva-Donnelly L, Szewczak AA, Gutell RR, Strobel SA. 1998. *RNA* 4: 498–519
48. Christian EL, Yarus M. 1993. *Biochemistry* 32:4475–80
49. Piccirilli JA, Vyle JS, Caruthers MH, Cech TR. 1993. *Nature* 361:85–88
50. Yoshida A, Sun S, Piccirilli JA. 1999. *Nat. Struct. Biol.* 6:318–21
51. Weinstein LB, Jones BCNM, Cosstick R, Cech TR. 1997. *Nature* 388:805–8
52. Harris ME, Pace NR. 1995. *RNA* 1:210–18
53. Nesbitt S, Hegg LA, Fedor MJ. 1997. *Chem. Biol.* 4:619–30
54. Young KJ, Gill F, Grasby JA. 1997. *Nucleic Acids Res.* 25:3760–66
55. Sood VD, Beattie TL, Collins RA. 1998. *J. Mol. Biol.* 282:741–50
56. Jeoung Y, Kumar PKR, Suh Y, Taira K, Nishikawa S. 1994. *Nucleic Acids Res.* 22:3722–27
57. Sontheimer EJ, Gordon PM, Piccirilli JA. 1999. *Genes Dev.* 13:1729–41
58. Pley HW, Flaherty KM, McKay DB. 1994. *Nature* 372:68–74
59. Scott WG, Finch JT, Klug A. 1995. *Cell* 81:991–1002
60. Ferre-D'Amare AR, Zhou K, Doudna JA. 1998. *Nature* 395:567–74
61. Hoogstraten CG, Legault P, Pardi A. 1998. *J. Mol. Biol.* 284:337–50
62. Wedekind JE, McKay DB. 1999. *Nat. Struct. Biol.* 6:261–68
63. Cai Z, Tinoco I Jr. 1996. *Biochemistry* 35:6026–36
64. Butcher SE, Allain FH-T, Feigon J. 1999. *Nat. Struct. Biol.* 6:212–16
65. Earnshaw DJ, Masquida B, Muller S, Sigurdsson ST, Eckstein F, et al. 1997. *J. Mol. Biol.* 274:197–212
66. Ryder SP, Strobel SA. 1999. *J. Mol. Biol.* 291:295–311
67. Uhlenbeck OC. 1987. *Nature* 328:596–600
68. Haseloff J, Gerlach WL. 1988. *Nature* 334:585–91
69. Tuschl T, Gohlke C, Jovin TM, Westhof E, Eckstein F. 1994. *Science* 266:785–89
70. Bassi GS, Mollegaard NE, Murchie AI, von Kitzing E, Lilley DM. 1995. *Nat. Struct. Biol.* 2:45–55
71. Jack A, Ladner JE, Rhodes D, Brown RS, Klug A. 1977. *J. Mol. Biol.* 111:315–28
72. van Tol H, Buzayan JM, Feldstein PA, Eckstein F, Bruening G. 1990. *Nucleic Acids Res.* 18:1971–75
73. Slim G, Gait MJ. 1991. *Nucleic Acids Res.* 19:1183–88
74. Dahm SC, Derrick WB, Uhlenbeck OC. 1993. *Biochemistry* 32:13040–45

75. Pontius BW, Lott WB, von Hippel PH. 1997. *Proc. Natl. Acad. Sci. USA* 94: 2290–94
76. Lott WB, Pontius BW, von Hippel PH. 1998. *Proc. Natl. Acad. Sci. USA* 95: 542–47
77. Murray JB, Seyhan AA, Walter NG, Burke JM, Scott WG. 1998. *Chem. Biol.* 5: 587–95
78. Scott WG, Murray JB, Arnold JRP, Stoddard BL, Klug A. 1996. *Science* 274: 2065–69
79. McKay DB. 1996. *RNA* 2:395–403
80. Feig AL, Scott WG, Uhlenbeck OC. 1998. *Science* 279:81–84
81. Murray JB, Terwey DP, Maloney L, Karpeisky A, Usman N, et al. 1998. *Cell* 92:665–73
82. Peracchi A, Beigelman L, Scott EC, Uhlenbeck OC, Herschlag D. 1997. *J. Biol. Chem.* 272:26822–26
83. Brown RS, Dewan JC, Klug A. 1985. *Biochemistry* 24:4785–801
84. Pan T, Uhlenbeck OC. 1992. *Nature* 358: 560–63
85. Pan T, Uhlenbeck OC. 1992. *Biochemistry* 31:3887–95
86. Pan T, Dichtl B, Uhlenbeck OC. 1994. *Biochemistry* 33:9561–65
87. Sugimoto N, Ohmichi T. 1996. *FEBS Lett.* 393:97–100
88. Ohmichi T, Sugimoto N. 1997. *Biochemistry* 36:3514–21
89. Legault P, Hoogstraten CG, Metlitzky E, Pardi A. 1998. *J. Mol. Biol.* 284:325–35
90. Thill G, Vasseur M, Tanner NK. 1993. *Biochemistry* 32:4254–62
91. Wu H, Lai MMC. 1990. *Mol. Cell. Biol.* 10:5575–79
92. Rosenstein SP, Been MD. 1990. *Biochemistry* 29:8011–16
93. Suh Y, Kumar PKR, Taira K, Nishikawa S. 1993. *FEBS Lett.* 21:3277–80
94. Ferre-D'Amare AR, Doudna JA. 2000. *J. Mol. Biol.* 295:541–56
95. Perrotta AT, Been MD. 1990. *Nucleic Acids Res.* 18:6821–27
96. Tanner NK, Schaff S, Thill G, Petit-Koskas E, Crain-Denoyville A, Westhof E. 1994. *Curr. Biol.* 4:488–98
97. Bravo C, Lescure F, Laugaa F, Fourrey J, Favre A. 1996. *Nucleic Acids Res.* 24: 1351–59
98. Huang Z, Ping Y, Wu H. 1997. *FEBS Lett.* 413:299–303
99. Wu H, Lin Y, Lin F, Makino S, Chang M, Lai MMC. 1989. *Proc. Natl. Acad. Sci. USA* 86:1831–35
100. Perrotta AT, Shih I-H, Been MD. 1999. *Science* 286:123–26
101. Shih I-H, Been MD. 1999. *RNA* 5:1140–48
102. Matysiak M, Wrzesinski J, Ciesiolka J. 1999. *J. Mol. Biol.* 291:283–94
103. Lafontaine DA, Ananvoranich S, Perreault J-P. 1999. *Nucleic Acids Res.* 27: 3236–43
104. Hampel A, Tritz R. 1989. *Biochemistry* 28:4929–33
105. Haseloff J, Gerlach W. 1989. *Gene* 82: 43–52
106. Feldstein P, Buzayan J, Bruening G. 1989. *Gene* 82:53–61
107. Walter NG, Hampel KJ, Brown KM, Burke JM. 1998. *EMBO J.* 17:2378–91
108. Hampel KJ, Walter NG, Burke JM. 1998. *Biochemistry* 37:14672–82
109. Hampel A, Cowan JA. 1997. *Chem. Biol.* 4:513–17
110. Grabowski PJ, Zaug AJ, Cech TR. 1981. *Cell* 23:467–76
111. Michel F, Westhof E. 1990. *J. Mol. Biol.* 216:585–610
112. Shan S, Narlikar GJ, Hershlag D. 1999. *Biochemistry* 38:10976–88
113. Heuer TS, Chandry PS, Belfort M, Celander DW, Cech TR. 1991. *Proc. Natl. Acad. Sci. USA* 88:11105–9
114. Michel F, Hanna M, Green R, Bartel DP, Szostak JW. 1989. *Nature* 342:391–95
115. Pyle AM, Murphy FL, Cech TR. 1992. *Nature* 358:123–28
116. Wang JF, Cech TR. 1992. *Science* 256: 526–29

117. Wang JF, Downs WD, Cech TR. 1993. *Science* 260:504–8
118. Strobel SA, Cech TR. 1993. *Biochemistry* 32:13593–604
119. Strobel SA, Cech TR. 1995. *Science* 267:675–79
120. Lehnert V, Jaeger L, Michel F, Westhof E. 1996. *Chem. Biol.* 3:993–1009
121. Tanner MA, Cech TR. 1995. *RNA* 1:349–50
122. Szewczak AA, Ortoleva-Donnelly L, Ryder SP, Moncoeur E, Strobel SA. 1998. *Nat. Struct. Biol.* 5:1037–42
123. Szewczak AA, Ortoleva-Donnelly L, Zivarts MV, Oyelere AK, Kazantsev AV, Strobel SA. 1999. *Proc. Natl. Acad. Sci. USA* 96:11183–88
124. Cate JH, Gooding AR, Podell E, Zhou K, Golden BL, et al. 1996. *Science* 273:1678–85
125. Golden BL, Gooding AR, Podell ER, Cech TR. 1998. *Science* 282:259–64
126. Strobel SA, Shetty K. 1997. *Proc. Natl. Acad. Sci. USA* 94:2903–8
127. Strobel SA, Ortoleva-Donnelly L, Ryder SP, Cate JH, Moncoeur E. 1998. *Nat. Struct. Biol.* 5:60–66
128. Puglisi JD, Wyatt JR, Tinoco I Jr. 1991. *Acc. Chem. Res.* 24:152–58
129. Mohr G, Caprara MG, Guo Q, Lambowitz AM. 1994. *Nature* 370:147–50
130. Joyce GF, van der Horst G, Inoue T. 1989. *Nucleic Acids Res.* 17:7879–89
131. Beaudry AA, Joyce GF. 1990. *Biochemistry* 29:6534–39
132. van der Horst G, Christian A, Inoue T. 1991. *Proc. Natl. Acad. Sci. USA* 88:184–88
133. Murphy FL, Cech TR. 1993. *Biochemistry* 32:5291–300
134. Laggerbauer B, Murphy FL, Cech TR. 1994. *EMBO J.* 13:2669–76
135. Doherty EA, Doudna JA. 1997. *Biochemistry* 36:3159–69
136. Doherty EA, Herschlag D, Doudna JA. 1999. *Biochemistry* 38:2982–90
137. Engelhardt MA, Doherty EA, Knitt DS, Doudna JA, Herschlag D. 2000. *Biochemistry* 39:2639–51
138. Michel F, Umesono K, Ozeki H. 1989. *Gene* 82:5–30
139. Haas ES, Brown JW. 1998. *Nucleic Acids Res.* 26:4093–99
140. Massire C, Jaeger L, Westhof E. 1998. *J. Mol. Biol.* 279:773–93
141. Weeks KM, Cech TR. 1995. *Cell* 82:221–30
142. Batey R, Williamson JR. 1998. *RNA* 4:984–97
143. Wimberly BT, Guymon R, McCutcheon JP, White SW, Ramakrishnan V. 1999. *Cell* 97:491–502
144. Conn G, Draper DE, Lattman EE, Gittis AG. 1999. *Science* 284:1171–74
145. Clemons WMJ, May JL, Wimberly BT, McCutcheon JP, Capel MS, Ramakrishnan V. 1999. *Nature* 400:833–40
146. Ban N, Nissen P, Hansen J, Capel M, Moore PB, Steitz TA. 1999. *Nature* 400:841–47
147. Cate JH, Yusupov MM, Yusupova GZ, Earnest TN, Noller HF. 1999. *Science* 285:2095–104