The chemical repertoire of natural ribozymes

Jennifer A. Doudna* & Thomas R. Cech†

**Department of Molecular and Cell Biology, and Howard Hughes Medical Institute, University of California, Berkeley, California 94720, USA (e-mail: doudna@uclink.berkeley.edu)*
†Howard Hughes Medical Institute, 4000 Jones Bridge Road, Chevy Chase, Maryland 20815, USA (e-mail: thomas.cech@colorado.edu)

Although RNA is generally thought to be a passive genetic blueprint, some RNA molecules, called ribozymes, have intrinsic enzyme-like activity — they can catalyse chemical reactions in the complete absence of protein cofactors. In addition to the well-known small ribozymes that cleave phosphodiester bonds, we now know that RNA catalysts probably effect a number of key cellular reactions. This versatility has lent credence to the idea that RNA molecules may have been central to the early stages of life on Earth.

How life began on Earth is one of the great scientific mysteries. Molecular biologists have long suspected that RNA molecules were key to the process, in part because RNA has essential roles in a most fundamental process — protein synthesis — within all cells. The first example of an RNA molecule that forms a catalytic active site for a series of precise biochemical reactions was reported 20 years ago: the self-splicing pre-ribosomal RNA (rRNA) of the ciliate *Tetrahymena*. Although there was only one example, the word 'ribozyme' was coined for the general concept of an RNA molecule with enzyme-like activity[1]. The following year catalytic activity was discovered in the RNA component of a ribonucleoprotein enzyme, ribonuclease (RNase) P, providing the first example of a multiple-turnover enzyme using RNA-based catalysis[2]. These findings lent increased credibility to the hypothesis of an RNA world, where RNA served both as the genetic material and the principal cellular enzyme, probably assisted in the latter role by metal ions, amino acids and other small-molecule cofactors. The RNA world hypothesis posits that as cellular metabolism became more sophisticated, increasing demands on biocatalysts provided the impetus for the transition to protein enzymes. Descendants from this proposed RNA-dominated era inhabit today's world in the form of naturally occurring ribozymes present in organisms ranging from bacteria to humans (Table 1).

Although the known natural cellular and viral ribozymes catalyse only phosphodiester transfer chemistry, ribozymes obtained through *in vitro* selection techniques can exhibit the sort of biochemical sophistication necessary to support cellular metabolism. Starting with a pool of random RNA sequences, molecules possessing a desired activity are isolated through successive cycles of activity selection, reverse transcription of the 'winners' into DNA and amplification of those sequences by the polymerase chain reaction. This methodology has allowed identification of ribozymes that form a nucleotide from a base plus a sugar[3], synthesize amide bonds[4,5], form Michael adducts such as those involved in the methylation of uridine monophosphate to give thymidine monophosphate[6], and form acyl-coenzyme A, which is found in many protein enzymes[7]. It is tantalizing to think that these ribozymes are analogues of missing links in a transition from an RNA world to contemporary biology (ref. 8, and see review in this issue by Joyce, pages 214–221). Because the structures and chemical mechanisms of *in vitro*-selected ribozymes are largely unknown at present, we focus here on the more extensively studied natural ribozymes.

RNA-based catalysis

How do RNA catalysts compare to their better-known protein enzyme counterparts? First, ribozyme rate enhancements can be substantial. For example, the rate constant for chemistry of the self-cleaving hepatitis delta virus (HDV) ribozyme is estimated at 10^2–10^4 s^{-1}, which is close to the maximal cleavage rate of RNase A (1.4×10^3 s^{-1} at 25 °C). The *Tetrahymena* group I self-splicing intron also has a rate constant for its chemical step that is comparable to those of protein enzymes. Second, ribozymes can use cofactors such as imidazole during catalysis[9,10], and they can be switched on and off by the binding of small-molecule allosteric effectors[11]. Finally, molecular structures have revealed that ribozymes, like protein catalysts, fold into specific three-dimensional shapes that can harbour deep grooves and solvent-inaccessible active sites. These tertiary structures facilitate catalysis in part by orienting substrates adjacent to catalytic groups and metal ions.

To facilitate chemical transformations, catalysts stabilize the transition state between substrate and product. Both protein and RNA catalysts may achieve this by adding or removing protons during a reaction, orienting substrates so that they are optimally positioned to react, and using binding interactions away from the reaction site to 'force' an unfavourable contact that is relieved in the transition state. Here we discuss the structural and chemical basis for RNA catalysis as determined for several of the naturally occurring ribozymes. The lack of diverse functional groups in RNA molecules and the propensity for RNA to bind metal ions led to early hypotheses that all ribozymes might act as metalloenzymes, positioning metal ions for direct roles in catalysis. This seems to be true for group I and group II self-splicing introns and RNase P, and in the *Tetrahymena* group I intron these metal ions and their specific functions have been identified. Surprisingly, however, those ribozymes that perform site-specific strand scission — the hammerhead, hairpin, *Neurospora* Varkud satellite (VS) and HDV ribozymes — may use diverse catalytic mechanisms, and none has emerged clearly as a metalloenzyme.

Site-specific RNA self-cleavage

The hammerhead, HDV, hairpin and VS ribozymes are small RNA structures of ~40–160 nucleotides that catalyse

site-specific self-cleavage (Table 1). Found in viral, virusoid or satellite RNAs, they process the multimeric products of rolling-circle replication into genome-length strands. Although the reaction catalysed by these ribozymes is the same as that of many protein RNases (Fig. 1a), they act only at specific phosphodiester bonds by using base-pairing and other interactions to align the cleavage site within the ribozyme active site. The evolutionary maintenance of these sequences may result from the relative simplicity and efficiency of RNA-catalysed RNA strand scission.

Hammerhead ribozyme

The hammerhead, at ~40 nucleotides, is the smallest of the naturally occurring ribozymes, and mediates rolling-circle replication within circular virus-like RNAs that infect plants. Recent experiments show that the hammerhead motif (Fig. 1b) is the most efficient self-cleaving sequence that can be isolated from randomized pools of RNA, suggesting that it may have arisen multiple times during the evolution of functional RNA molecules[12]. Consisting of three short helices connected at a conserved sequence junction, the hammerhead catalyses site-specific cleavage of one of its own phosphodiester bonds via nucleophilic attack of the adjacent 2′-oxygen at the scissile phosphate (Fig. 1a).

The simplicity of the hammerhead secondary structure lent itself to the design of two-piece constructs in which the strand containing the cleavage site was separated from the rest of the self-cleaving RNA. By treating one strand as the substrate and the other as the enzyme, multiple-turnover cleavage occurred with a typical rate of 1 molecule per minute at physiological salt concentrations, consistent with a substantial 10^9-fold rate enhancement over the uncatalysed rate of nonspecific RNA hydrolysis. Initial studies revealed a requirement for a divalent metal ion for catalysis, leading to the idea that site-specific positioning of a metal ion such as magnesium might enable efficient deprotonation of the attacking 2′-hydroxyl nucleophile. Later studies at much higher ionic strength (4-M monovalent salt) showed that the hammerhead as well as the hairpin and VS ribozymes could react nearly as fast in the absence of divalent ions[13]. This discovery suggested two distinct possibilities: either these ribozymes use a different catalytic mechanism in the presence of high, non-physiological concentrations of monovalent salts, or the divalent metal ion requirement at low salt concentrations serves a structural rather than a purely chemical function.

The unveiling of the crystal structure of the hammerhead ribozyme in 1994, the first ribozyme structure to be determined, revealed a Y-shaped conformation in which nucleotides essential for catalysis were clustered at the junction of the three helical arms[14] (Fig. 1c). Since then, additional crystal structures of the hammerhead ribozyme have provided 'snapshots' of the RNA at several steps along the catalytic reaction pathway: the initial pre-cleaved state, two sequential conformational changes that precede cleavage and rotate the scissile phosphate to be in line with the attacking 2′-oxygen nucleophile, and the post-cleavage product state[15–19]. Together these structures have led to a model of ribozyme cleavage involving precise positioning of the reactive groups by the structure of the ribozyme. However, it has been difficult to ascertain from these structures what the role of bound divalent metal ions might be. Although several divalent ions were identified unambiguously in the crystal structures, they were not situated close enough to the site of catalysis to support a direct role in RNA cleavage.

Site-specific substitution of phosphate oxygens with sulphur atoms, which is readily achieved by solid-phase synthetic methods, enabled direct analysis of the effects of disrupting divalent metal-ion binding sites that were potentially involved in catalysis[20]. These investigations led to evidence for the direct simultaneous coordination of a single metal ion by the scissile phosphate and a second phosphate oxygen located 20 Å away in the crystal structure[21,22]. It was proposed that the crystal structures might represent the 'ground state' conformation of the hammerhead ribozyme, and that prior to catalysis the RNA conformation changed significantly but transiently to bring the critical catalytic metal ion proximal to the cleavage site. More recently, molecular modelling and kinetic analysis of the hammerhead cleavage reaction in the presence of monovalent versus divalent salts support the idea that divalent metal ions are not essential to the catalytic step, but instead stabilize the active ribozyme structure[19,23–25].

Whether hammerhead catalysis requires a global conformational change or merely a local rearrangement is not yet resolved. In either case, orientation of reactants within the ribozyme active site

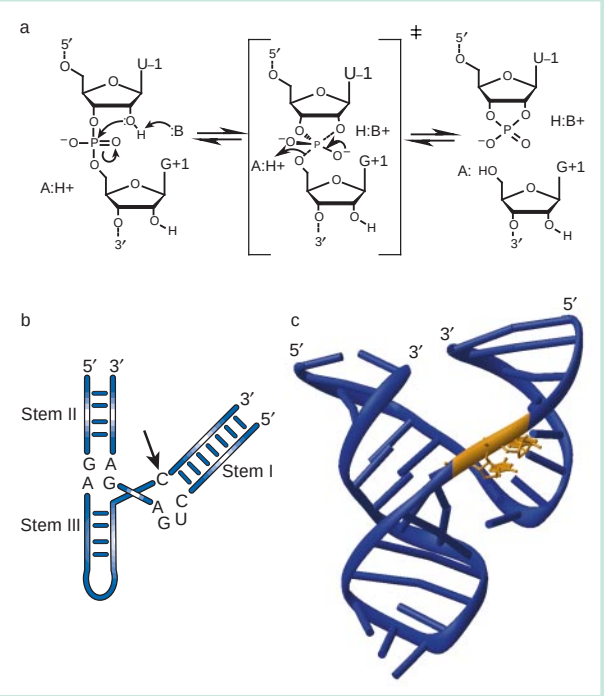

Figure 1 Mechanism of RNA-catalysed self-cleavage. **a**, General mechanism of ribonucleases and small self-cleaving ribozymes. The 2′-hydroxyl adjacent to the scissile phosphate is activated for nucleophilic attack by abstraction of its proton. Concurrently, a proton is donated to stabilize the developing negative charge on the leaving group oxygen. **b**, Secondary structure of the hammerhead ribozyme. Nucleotides important for catalytic activity are indicated; the cleavage site is indicated by an arrow. **c**, Crystal structure of the hammerhead ribozyme. Coordinates are from ref. 14. The nucleotides flanking the scissile bond are shown in gold.

Table 1 Naturally occurring ribozymes and ribonucleoprotein enzymes

Ribozyme	Sequenced examples	Size (nt)	Activity (reaction product)
Hammerhead	11	40	Self-cleavage via transesterification (2′,3′ cyclic phosphate)
Hepatitis delta virus	2	90	
Hairpin	1	70	
Varkud satellite	1	160	
Group I intron	>1,500	210	Self-splicing via transesterification (3′-OH)
Group II intron	>700	500	
RNase P*	>500	300	Pre-tRNA processing via hydrolysis (3′-OH)
Spliceosome* (U2+U6 snRNAs)	70,50	180,100	RNA splicing via transesterification (3′-OH)
Ribosome* (23S rRNA)	>900	2,600	Peptidyl transfer (amide)

Number of sequenced examples is a snapshot as of 2002 and is influenced by DNA-sequencing strategies and database upkeep; it may provide a rough indication of relative abundance. RNAs in any group vary in size; the size provided here indicates the lower end of the length distribution for the natural examples. See www.rna.icmb.utexas.edu and www.jwbrown.mbio.ncsu.edu/RnaseP/.
*Ribonucleoprotein enzymes. RNase P: bacterial and archaebacterial RNAs have the relevant activity in the absence of protein. Spliceosome: U2 and U6 small nuclear RNAs (snRNAs) alone show an activity related to the natural activity. Ribosome: no activity has yet been observed with protein-free, large-subunit rRNA.

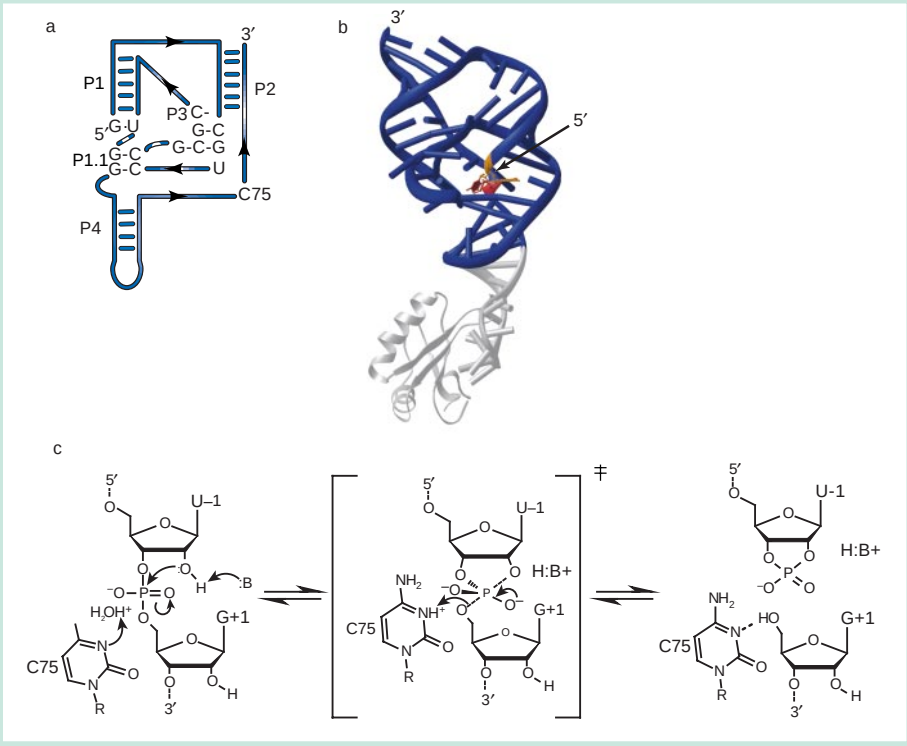

Figure 2 Structure of the hepatitis delta virus (HDV) ribozyme. **a**, Secondary structure of the genomic form of the HDV ribozyme; the cytosine residue essential to catalytic activity (C75) is indicated, and the cleavage site is marked by the 5′ end of the RNA strand. **b**, Crystal structure of the product form of the HDV ribozyme. The active-site cytosine is shown in red, the 5′ nucleotide of the ribozyme in gold, and the U1a RNA-binding-domain protein and its cognate RNA-binding site in grey (this has been engineered into the construct to assist crystallization). **c**, Proposed mechanism of general acid catalysis by C75, in which the protonated form of the C donates a proton to the leaving group during catalysis (compare with Fig. 1a).

probably contributes significantly to the rate of site-specific strand scission in the hammerhead. This is achieved through the unique structure of this RNA motif, conferred by the secondary structure and the presence of multiple conserved nucleotides in the active site. Whether these nucleotides provide anything else, such as general acid–base catalysis, is unknown.

Hepatitis delta virus and hairpin ribozymes

The HDV and hairpin ribozymes catalyse the same chemical reaction as that of the hammerhead, and they are likewise responsible for cleaving intermediates generated during rolling-circle replication of the HDV and a plant virus satellite RNA, respectively. Crystal structures of these ribozymes showed that in each case the RNA forms an enclosed cleft in which strand scission takes place[26,27]. Furthermore, neither ribozyme seems to coordinate a divalent metal ion at the site of catalysis, but instead positions functionally essential nucleotides proximal to the substrate in a configuration suggesting the possibility of their direct role in catalysis.

The potential for RNA to use general acid–base chemistry during catalysis was unexpected. Functional groups within proteins that have pK_a values near neutrality, and thus can donate or accept a proton readily under physiological conditions, can act as general acids or general bases to shuttle protons during enzyme catalysis. But the lack of functional groups within RNA with pK_a values near physiological pH (6–7) means that for RNA to function in this way, one or more of its functional groups must have a pK_a significantly shifted towards neutrality. In RNA, adenine and cytosine have the potential for protonation of their ring nitrogens, N1 and N3, respectively, but the pK_a values for the free nucleosides are 3.5 and 4.2. However, A and C residues with substantially shifted pK_a values have been detected in small functional RNAs, presumably owing to the structural environment of the nucleotide[28–32].

In the HDV ribozyme, a cytosine base essential to catalytic function is positioned in a cleft adjacent to the site of cleavage in the RNA (Fig. 2a, b). A network of potential hydrogen bonds to this nucleotide is consistent with stabilization of a protonated form of the cytosine that might allow it to donate or accept a proton at some stage during catalysis[26]. This feature could be useful for mediating catalysis by pulling a proton off the attacking 2′-oxygen nucleophile, or by providing a proton to the 5′-oxygen leaving group (Fig. 1a). But does this in fact occur? Although detecting the movement of protons within an enzyme, or ribozyme, active site is difficult to achieve directly, several clever experiments have provided indirect evidence that such a mechanism is at work in the HDV ribozyme.

In one set of experiments, catalytic activity of ribozymes with point mutations at the critical cytosine was partially restored in the presence of imidazole, the side chain of histidine that readily accepts or donates a proton in many protein enzyme active sites[9]. Furthermore, the measured pK_a values of a series of restored reactions correlated with those of the imidazole analogues that promoted cleavage in the mutants[33]. In a second series of experiments, kinetic isotope effects and correlation of reaction pK_a with the pK_a of different bases placed at the position of the critical C supported a direct role of this nucleotide in proton shuttling during catalysis[34,35]. However, the pK_a of the active-site C is apparently transiently shifted during the reaction, as the shift is not detected within the product or precursor forms of the ribozyme[36]. Together, the current data support a model in which the C acts as a general acid during the reaction to donate a proton to the 5′-bridging oxygen (Fig. 2c). A hydrated metal ion coordinated near the ribozyme active site may abstract a proton from the 2′-hydroxyl nucleophile[37] (Fig. 1a).

In the hairpin ribozyme the situation is less clear. Although the crystal structure of a precursor form of the RNA suggested that an active-site adenosine might adopt the role of a general acid or base[27] (Fig. 3a, b), nucleotide analogue interference experiments have not provided evidence for a shifted pK_a at this position[38]. However, the lack of a requirement for divalent metal ions during hairpin ribozyme cleavage implies that a metal ion-independent mechanism is at work[39–41].

Self-splicing introns

Group I introns

Group I introns have been found to interrupt genes for rRNA, transfer RNA (tRNA) and messenger RNA (mRNA) in far-reaching

corners of biology, including the nuclei of protozoa, the mitochondria of fungi, the chloroplasts of algae, and bacteria and their phages. They are defined as a group by their common core secondary structure, consisting of an array of nine base-paired elements (P1–P9), and by their common mechanism of self-splicing. They accomplish splicing by a two-step transesterification mechanism initiated by an exogenous molecule of guanosine or guanosine triphosphate (GTP; Fig. 4a).

Because the excised intron still contains the active site for transesterification, it can be re-engineered to give a true catalyst that can cleave or ligate exogenous substrate molecules intermolecularly ('*in trans*'). The *Tetrahymena* version of this RNA enzyme, together with oligonucleotide substrates that can be synthesized with subtle chemical variations, has facilitated the dissection of the reaction pathway into its elemental steps. First, the RNA substrate base-pairs to an internal guide sequence within the ribozyme, forming the P1 helix. Second, specific ribose 2′-hydroxyl groups along the minor groove of the P1 helix promote its docking into the active site. Third, guanosine (or a 5′-phosphorylated analogue such as GTP) binds to the G site within P7. Fourth, the 3′-hydroxyl of the G acts as a nucleophile, cleaving the 5′ splice-site phosphate with inversion of configuration. Finally, products are released; the slow release of the product bound by base pairing plus tertiary interactions is rate limiting for multiple-turnover cleavage under typical conditions[42].

Many protein enzymes that promote phosphoryl transfer reactions use metal ions for catalysis, and group I ribozymes use the same trick[43]. The number and location of the active-site metal ions has been investigated by substituting individual phosphate or ribose oxygen atoms with sulphur or with an amino group, and then testing for changes in metal-ion specificity (a procedure known as 'thiophilic metal-ion rescue'). The most highly supported current model is shown in Fig. 4b. One metal ion (M_B) helps deprotonate the 3′-oxygen of the G nucleophile, while another (M_A) stabilizes the developing negative charge on the leaving-group oxygen in the transition state. M_C could assist in positioning the substrates with respect to one another and, along with M_A, could help stabilize the trigonal, bipyramidal transition state[44].

The enzyme mechanism of the group I introns — involving a nucleotide-binding site, nucleophilic attack and metal-ion catalysis — is commonplace in the world of protein enzymes. But clearly the ribozyme active site must be constructed differently, given the charged and hydrophilic nature of the nucleic acid building blocks and the limited diversity of their side chains compared to amino acids. So how might a catalytic active site be built out of ribonucleotides?

The first detailed view of active-site construction came with the crystal structure of the 160-nucleotide P4–P6 domain of the *Tetrahymena* intron, an attractive target because it folds into the same structure as an excised domain as it does in the context of the whole ribozyme. This structure revealed how long-range base triples and divalent cation-mediated structures can fold an RNA molecule into a globular structure with an interior that is relatively inaccessible to solvent[45,46]. The way in which this domain and the G-site-containing domain combine to create a concave active site was seen at modest resolution in the crystal structure of a 240-nucleotide active ribozyme[47], a structure whose general features had been predicted by modelling based on comparative phylogenetic analysis[48]. Finally, the way the structure embraces the P1 substrate helix was modelled by mapping sites of chemical modification that perturb the reaction[49]. Future goals are to obtain a high-resolution crystal structure of an entire group I ribozyme with bound substrates, and to locate the proposed three catalytic metals within that structure.

Group II introns

Group II introns, found in bacteria and in organellar genes of eukaryotic cells, catalyse precise self-excision and ligation of the flanking RNA sequences to form a mature transcript. In a mechanism distinct from that of the group I introns, the group II reaction involves nucleophilic attack by the 2′-hydroxyl of a specific adenosine within the intron — the 'branch site' — to form a branched or lariat-type structure (Fig. 4a). Magnesium ions coordinated within the intron are thought to have a direct role in catalysis[50–52], and several studies have revealed aspects of the intron tertiary structure that are essential to catalytic function[53–56]. Models of group II intron architecture have been proposed based on chemical probing, phylogenetic covariation and mutagenesis results[57,58]. Interestingly, some group II introns encode proteins that assist RNA splicing and can also enable efficient integration of the intron RNA into double-stranded DNA by reverse splicing and reverse transcription[59]. This reverse splicing activity promotes intron mobility by enabling insertion into targeted genes.

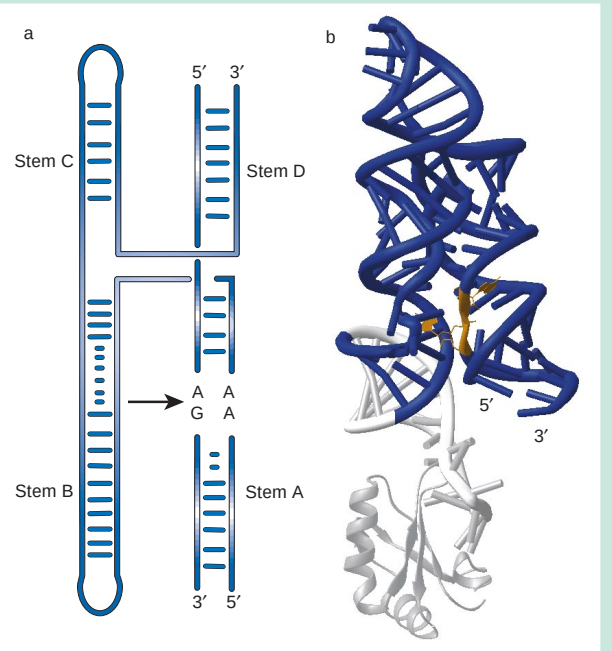

Figure 3 Structure of the hairpin ribozyme. **a**, Secondary structure of the hairpin ribozyme; conserved, functionally important nucleotides are shown explicitly. Dots indicate non-canonical base pairings. **b**, Crystal structure of a precursor form of the hairpin ribozyme. Nucleotides flanking the scissile bond are shown in gold, whereas the grey structure is the U1a RNA-binding-domain protein and its cognate RNA-binding site, engineered into the construct to assist crystallization.

Ribonuclease P

RNase P, found in all cells, catalyses site-specific hydrolysis of precursor RNA substrates including tRNA, 5S rRNA and the signal recognition particle RNA[60,61]. These substrates probably share structural features that enable efficient recognition by the RNase P substrate-binding site, positioning the reactive phosphate in each case for nucleophilic attack by a coordinated water molecule. The ribozyme is thought to be a metalloenzyme, a hypothesis supported by data from phosphorothioate substitution at the scissile phosphate of a substrate pre-tRNA[62,63].

RNase P is in fact an RNA–protein complex whose activity, at least in bacteria, resides within the RNA component. Furthermore, it is a true catalyst in the sense that each ribozyme complex catalyses the cleavage of multiple substrate RNAs. As with group I and II introns, much of the information about the secondary and tertiary structure of RNase P RNA has come from extensive phylogenetic covariation analysis of related sequences. Typically 300–400 nucleotides in length, it comprises two domains containing the substrate-recognition site and the ribozyme active site, respectively. Structural models

of a bacterial RNase P RNA have been proposed[64,65], but a crystal structure is not yet available. In human cells, the RNase P complex is larger and includes multiple protein components in addition to the RNA[66–68]. The human RNase P RNA is not catalytically active in the absence of protein, which has made it challenging to determine whether its active site is composed of RNA, protein or some combination of the two.

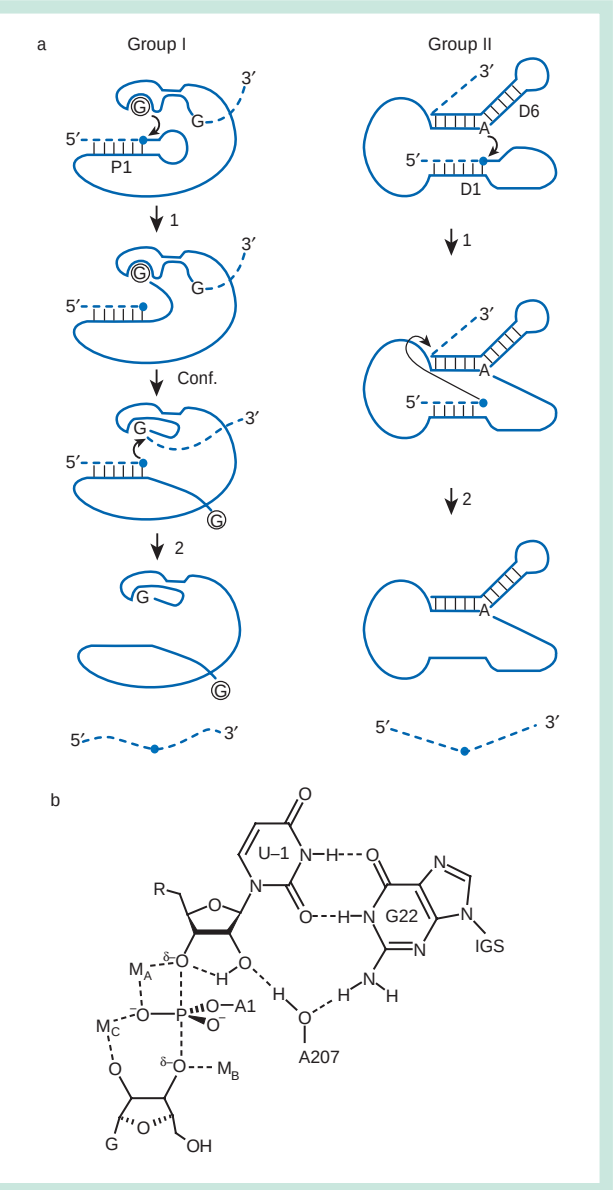

Figure 4 Self-splicing intron mechanisms. **a**, Pathways for group I and II intron self-splicing, with exons shown as dashed lines and introns as solid lines. For group I introns, step 1 shows how an intron-bound guanosine or GTP (circled) cleaves the 5′ splice site while becoming covalently attached to the 5′ end of the intron. 'Conf.' indicates a conformational change whereby the G at the 3′ end of the intron replaces the original G in the G-binding site. In step 2, the cleaved 5′ exon, still held to the intron by base pairing (P1), then cleaves the 3′ splice site; as a result, the exons are ligated and the intron excised. For group II introns, step 1 shows how an adenosine 2′-hydroxyl within domain 6 (D6) attacks the 5′ splice site, which is identified by base-pairing interactions involving domain 1 (D1); this results in a branched 'lariat' RNA intermediate. In step 2, the cleaved 5′ exon then attacks the 3′ splice site, ligating the exons and excising the lariat intron. **b**, Three-metal-ion mechanism for RNA cleavage catalysed by the *Tetrahymena* group I intron (adapted from ref. 44). The step shown is the same as step 1 in **a**; the cleavage site phosphate (between U–1 and A1) is recognized in part by interactions with G22 in the internal guide sequence (IGS).

Ribozyme activity, folding and dynamics

Protein enzymes that catalyse nucleophilic attack at a phosphate within RNA or a ribonucleotide apparently utilize different chemical mechanisms depending on the enzyme. For example, mammalian adenylyl cyclases function by a two-metal-ion mechanism[69], RNase A uses two histidines for general acid–bases catalysis[70], and the anthrax adenylyl cyclase exotoxin uses one histidine and a coordinated metal ion to activate the attacking nucleophile and stabilize the leaving group, respectively[71]. The fact that ribozymes also catalyse phosphodiester bond cleavage by a variety of mechanisms shows that RNA has a breadth of catalytic potential similar to protein enzymes.

Furthermore, like protein enzymes, ribozymes must fold into specific three-dimensional structures to function catalytically. How do these RNAs reach their active conformations? This problem has been the subject of intense study using *in vitro* systems, and several themes are emerging. RNA structures generally fold via a cooperative, hierarchical pathway in which tertiary interactions follow the formation of a stable secondary structure. Rates of tertiary structure formation vary from tens of milliseconds to several minutes, and for the large ribozymes, can be dominated by folding intermediates that transiently trap the RNA in non-native or partially folded conformations[72–75]. Single-molecule experiments, in which individual ribozyme molecules are analysed by fluorescence microscopy or mechanical tethering, reveal that multiple folding pathways can exist for a particular RNA sequence[76–78]. It is not yet clear how these observations relate to RNA folding processes *in vivo*, because proteins may assist ribozyme assembly either through direct RNA binding[79,80] or through covalent modification of specific nucleotides. In addition, RNA tertiary structures and even secondary structures can undergo rapid conformational changes[81–85]. Such dynamics may be essential for progress through a catalytic cycle. On the other hand, too much structural flexibility may hamper catalysis, providing the incentive for evolution of ribonucleoprotein enzymes rather than pure ribozymes.

A starring role for ribozymes?

If the RNA world had a lengthy head-start over the protein catalyst world, why are RNA catalysts relatively minor players in modern cells? In fact, they may be much more central to cell biology than was previously believed. The ribosome, which is responsible for information-directed protein synthesis in all of life, is composed of three (or in some cases four) RNA molecules along with several dozen proteins. No protein subunit has ever been identified as a peptidyl transferase enzyme, and for more than 20 years, evidence for a primary role of RNA in this activity has accumulated[86]. The most direct evidence came with the deduction of the crystal structure of the large subunit[87], in which the peptidyl transferase centre was precisely located by binding a small-molecule inhibitor that is an analogue of the anionic tetrahedral intermediate in amide bond formation[88]. Remarkably, only RNA and no protein lies in the vicinity of the reaction centre, so the catalysis must be ribozymic. The authors suggested one possible mechanism involving a conserved adenine acting as a general base to abstract an amino proton from the amino acid[89], but subsequent mutagenesis of the key A has not provided strong support[90,91]. Identifying the rRNA's catalytic strategy is an important direction for future research.

Given the large size of the rRNA (Table 1), scientists are using *in vitro* evolution to find smaller peptidyl transferase ribozymes that might model the biological reaction[92]. Indeed, RNA catalysts have also been identified that accomplish two other steps of the protein synthesis pathway: formation of activated amino acid adenylates[93] and transfer of an amino acid to the 3′-oxygen of a tRNA-like acceptor[94,95].

Another ribonucleoprotein catalyst found in eukaryotic cells is the spliceosome, which assembles with nuclear pre-mRNAs and splices out the major class of introns (which are not self-splicing). For many years, a popular hypothesis has been that the spliceosome is also a molecular fossil from the RNA world — with several enzymatic RNAs acting intermolecularly via a mechanism analogous to that used intramolecularly by self-splicing group II introns[96,97]. As with the ribosome, most of the evidence for this hypothesis has been circumstantial until very recently.

Now Valadkhan and Manley[98] have shown that two spliceosomal RNAs, the U2 and U6 small nuclear RNAs, can bind an RNA substrate containing the sequence of the intron branch site and promote a splicing-related reaction in the absence of any of the numerous spliceosomal proteins. The reaction product is not the natural 'branch' — consisting of a nucleotide forming both 2′-5′ and 3′-5′ phosphodiester bonds — but instead a new product consistent with a phosphotriester. (This is surprising from a chemical perspective, because it would require hydroxyl as the leaving group from a pentavalent phosphorous intermediate or transition state.) This new RNA-catalysed reaction will undoubtedly stimulate fresh investigations of the mechanism by which spliceosomal RNAs catalyse mRNA splicing, and adds weight to the proposition that remnants of the RNA world are still among us.

Future directions

As some of the first RNAs to be studied in structural and mechanistic detail, ribozymes have provided many important insights into RNA function at a fundamental chemical level. Although much progress has been made, many interesting questions remain to be addressed. Determining detailed reaction mechanisms for RNA catalysts, including large RNA–protein complexes such as the ribosome and the spliceosome, will be a priority, as well as exploring the chemical mechanisms of ribozymes identified by *in vitro* selection. In addition to revealing new aspects of RNA biology, these investigations may shed light on aspects of the proposed RNA world and the role of RNA in early evolution. □

1. Kruger, K. *et al.* Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* **31,** 147–157 (1982).
2. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35,** 849–857 (1983).
3. Unrau, P. J. & Bartel, D. P. RNA-catalysed nucleotide synthesis. *Nature* **395,** 260–263 (1998).
4. Lohse, P. A. & Szostak, J. W. Ribozyme-catalysed amino-acid transfer reactions. *Nature* **381,** 442–444 (1996).
5. Wiegand, T. W., Janssen, R. C. & Eaton, B. E. Selection of RNA amide synthases. *Chem. Biol.* **4,** 675–683 (1997).
6. Sengle, G., Eisenfuhr, A., Arora, P. S., Nowick, J. S. & Famulok, M. Novel RNA catalysts for the Michael reaction. *Chem. Biol.* **8,** 459–473 (2001).
7. Jadhav, V. R. & Yarus, M. Acyl-CoAs from coenzyme ribozymes. *Biochemistry* **41,** 723–729 (2002).
8. Wilson, D. S. & Szostak, J. W. *In vitro* selection of functional nucleic acids. *Annu. Rev. Biochem.* **68,** 611–647 (1999).
9. Perrotta, A. T., Shih, I. & Been, M. D. Imidazole rescue of a cytosine mutation in a self-cleaving ribozyme. *Science* **286,** 123–126 (1999).
10. Santoro, S. W., Joyce, G. F., Sakthivel, K., Gramatikova, S. & Barbas, C. F. III RNA cleavage by a DNA enzyme with extended chemical functionality. *J. Am. Chem. Soc.* **122,** 2433–2439 (2000).
11. Tang, J. & Breaker, R. R. Rational design of allosteric ribozymes. *Chem. Biol.* **4,** 453–459 (1997).
12. Salehi-Ashtiani, K. & Szostak, J. W. *In vitro* evolution suggests multiple origins for the hammerhead ribozyme. *Nature* **414,** 82–84 (2001).
13. Murray, J. B., Seyhan, A. A., Walter, N. G., Burke, J. M. & Scott, W. G. The hammerhead, hairpin and VS ribozymes are catalytically proficient in monovalent cations alone. *Chem. Biol.* **5,** 587–595 (1998).
14. Pley, H. W., Flaherty, K. M. & McKay, D. B. Three-dimensional structure of a hammerhead ribozyme. *Nature* **372,** 68–74 (1994).
15. Scott, W. G., Finch, J. T. & Klug, A. The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. *Cell* **81,** 991–1002 (1995).
16. Scott, W. G., Murray, J. B., Arnold, J. R., Stoddard, B. L. & Klug, A. Capturing the structure of a catalytic RNA intermediate: the hammerhead ribozyme. *Science* **274,** 2065–2069 (1996).
17. Murray, J. B. *et al.* The structural basis of hammerhead ribozyme self-cleavage. *Cell* **92,** 665–673 (1998).
18. Murray, J. B., Szoke, H., Szoke, A. & Scott, W. G. Capture and visualization of a catalytic RNA enzyme-product complex using crystal lattice trapping and X-ray holographic reconstruction. *Mol. Cell* **5,** 279–287 (2000).
19. Murray, J. B., Dunham, C. M. & Scott, W. G. A pH-dependent conformational change, rather than the chemical step, appears to be rate-limiting in the hammerhead ribozyme cleavage reaction. *J. Mol. Biol.* **315,** 121–130 (2002).
20. Scott, E. C. & Uhlenbeck, O. C. A re-investigation of the thio effect at the hammerhead cleavage site. *Nucleic Acids Res.* **27,** 479–484 (1999).
21. Peracchi, A., Beigelman, L., Scott, E. C., Uhlenbeck, O. C. & Herschlag, D. Involvement of a specific metal ion in the transition of the hammerhead ribozyme to its catalytic conformation. *J. Biol. Chem.* **272,** 26822–26826 (1997).
22. Wang, S., Karbstein, K., Peracchi, A., Beigelman, L. & Herschlag, D. Identification of the hammerhead ribozyme metal ion binding site responsible for rescue of the deleterious effect of a cleavage site phosphorothioate. *Biochemistry* **38,** 14363–14378 (1999).
23. Murray, J. B. & Scott, W. G. Does a single metal ion bridge the A-9 and scissile phosphate groups in the catalytically active hammerhead ribozyme structure? *J. Mol. Biol.* **296,** 33–41 (2000).
24. O'Rear, J. L. *et al.* Comparison of the hammerhead cleavage reactions stimulated by monovalent and divalent cations. *RNA* **7,** 537–545 (2001).
25. Curtis, E. A. & Bartel, D. P. The hammerhead cleavage reaction in monovalent cations. *RNA* **7,** 546–552 (2001).
26. Ferre-D'Amare, A. R., Zhou, K. & Doudna, J. A. Crystal structure of a hepatitis delta virus ribozyme. *Nature* **395,** 567–574 (1998).
27. Rupert, P. B. & Ferre-D'Amare, A. R. Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature* **410,** 780–786 (2001).
28. Rajagopal, P. & Feigon, J. Triple-strand formation in the homopurine:homopyrimidine DNA oligonucleotides d(G-A)4 and d(T-C)4. *Nature* **339,** 637–640 (1989).
29. Sklenar, V. & Feigon, J. Formation of a stable triplex from a single DNA strand. *Nature* **345,** 836–838 (1990).
30. Connell, G. J. & Yarus, M. RNAs with dual specificity and dual RNAs with similar specificity. *Science* **264,** 1137–1141 (1994).
31. Legault, P. & Pardi, A. In situ probing of adenine protonation in RNA by 13C NMR. *J. Am. Chem. Soc.* **116,** 8390–8391 (1994).
32. Ravindranathan, S., Butcher, S. E. & Feigon, J. Adenine protonation in domain B of the hairpin ribozyme. *Biochemistry* **39,** 16026–16032 (2000).
33. Shih, I. H. & Been, M. D. Involvement of a cytosine side chain in proton transfer in the rate-determining step of ribozyme self-cleavage. *Proc. Natl Acad. Sci. USA* **98,** 1489–1494 (2001).
34. Nakano, S., Chadalavada, D. M. & Bevilacqua, P. C. General acid-base catalysis in the mechanism of a hepatitis delta virus ribozyme. *Science* **287,** 1493–1497 (2000).
35. Nakano, S. & Bevilacqua, P. C. Proton inventory of the genomic HDV ribozyme in Mg^{2+}-containing solutions. *J. Am. Chem. Soc.* **123,** 11333–11334 (2001).
36. Luptak, A., Ferre-D'Amare, A. R., Zhou, K., Zilm, K. W. & Doudna, J. A. Direct pK_a measurement of the active-site cytosine in a genomic hepatitis delta virus ribozyme. *J. Am. Chem. Soc.* **123,** 8447–8452 (2001).
37. Nakano, S., Proctor, D. J. & Bevilacqua, P. C. Mechanistic characterization of the HDV genomic ribozyme: assessing the catalytic and structural contributions of divalent metal ions within a multichannel reaction mechanism. *Biochemistry* **40,** 12022–12038 (2001).
38. Ryder, S. P. *et al.* Investigation of adenosine base ionization in the hairpin ribozyme by nucleotide analog interference mapping. *RNA* **7,** 1454–1463 (2001).
39. Hampel, A. & Cowan, J. A. A unique mechanism for RNA catalysis: the role of metal cofactors in hairpin ribozyme cleavage. *Chem. Biol.* **4,** 513–517 (1997).
40. Nesbitt, S., Hegg, L. A. & Fedor, M. J. An unusual pH-independent and metal-ion-independent mechanism for hairpin ribozyme catalysis. *Chem. Biol.* **4,** 619–630 (1997).
41. Walter, N. G. & Burke, J. M. The hairpin ribozyme: structure, assembly and catalysis. *Curr. Opin. Chem. Biol.* **2,** 303 (1998).
42. Cech, T. R. & Herschlag, D. (eds) *Group I Ribozymes: Substrate Recognition, Catalytic Strategies and Comparative Mechanistic Analysis* (Springer, Berlin, 1996).
43. Narlikar, G. J. & Herschlag, D. Mechanistic aspects of enzymatic catalysis: lessons from comparison of RNA and protein enzymes. *Annu. Rev. Biochem.* **66,** 19–59 (1997).
44. Shan, S., Kravchuk, A. V., Piccirilli, J. A. & Herschlag, D. Defining the catalytic metal ion interactions in the *Tetrahymena* ribozyme reaction. *Biochemistry* **40,** 5161–5171 (2001).
45. Cate, J. H. *et al.* Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science* **273,** 1678–1685 (1996).
46. Juneau, K., Podell, E., Harrington, D. J. & Cech, T. R. Structural basis of the enhanced stability of a mutant ribozyme domain and a detailed view of RNA–solvent interactions. *Structure (Camb.)* **9,** 221–231 (2001).
47. Golden, B. L., Gooding, A. R., Podell, E. R. & Cech, T. R. A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme. *Science* **282,** 259–264 (1998).
48. Michel, F. & Westhof, E. Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.* **216,** 585–610 (1990).
49. Szewczak, A. A. *et al.* An important base triple anchors the substrate helix recognition surface within the *Tetrahymena* ribozyme active site. *Proc. Natl Acad. Sci. USA* **96,** 11183–11188 (1999).
50. Gordon, P. M., Sontheimer, E. J. & Piccirilli, J. A. Kinetic characterization of the second step of group II intron splicing: role of metal ions and the cleavage site 2′-OH in catalysis. *Biochemistry* **39,** 12939–12952 (2000).
51. Sigel, R. K., Vaidya, A. & Pyle, A. M. Metal ion binding sites in a group II intron core. *Nature Struct. Biol.* **7,** 1111–1116 (2000).
52. Gordon, P. M. & Piccirilli, J. A. Metal ion coordination by the AGC triad in domain 5 contributes to group II intron catalysis. *Nature Struct. Biol.* **8,** 893–898 (2001).
53. Jestin, J. L., Deme, E. & Jacquier, A. Identification of structural elements critical for inter-domain interactions in a group II self-splicing intron. *EMBO J.* **16,** 2945–2954 (1997).
54. Boudvillain, M., de Lencastre, A. & Pyle, A. M. A tertiary interaction that links active-site domains to the 5′ splice site of a group II intron. *Nature* **406,** 315–318 (2000).
55. Chu, V. T., Adamidi, C., Liu, Q., Perlman, P. S. & Pyle, A. M. Control of branch-site choice by a group II intron. *EMBO J.* **20,** 6866–6876 (2001).
56. Zhang, L. & Doudna, J. A. Structural insights into group II intron catalysis and branch-site selection. *Science* **295,** 2084–2088 (2002).
57. Costa, M., Michel, F. & Westhof, E. A three-dimensional perspective on exon binding by a group II self-splicing intron. *EMBO J.* **19,** 5007–5018 (2000).
58. Swisher, J., Duarte, C. M., Su, L. J. & Pyle, A. M. Visualizing the solvent-inaccessible core of a group II intron ribozyme. *EMBO J.* **20,** 2051–2061 (2001).
59. Yang, J., Zimmerly, S., Perlman, P. S. & Lambowitz, A. M. Efficient integration of an intron RNA into double-stranded DNA by reverse splicing. *Nature* **381,** 332–335 (1996).
60. Frank, D. N. & Pace, N. R. Ribonuclease P: unity and diversity in a tRNA processing ribozyme. *Annu. Rev. Biochem.* **67,** 153–180 (1998).

61. Morl, M. & Marchfelder, A. The final cut. The importance of tRNA 3′-processing. *EMBO Rep.* **2,** 17–20 (2001).
62. Warnecke, J. M., Held, R., Busch, S. & Hartmann, R. K. Role of metal ions in the hydrolysis reaction catalyzed by RNase P RNA from *Bacillus subtilis*. *J. Mol. Biol.* **290,** 433–445 (1999).
63. Warnecke, J. M., Sontheimer, E. J., Piccirilli, J. A. & Hartmann, R. K. Active site constraints in the hydrolysis reaction catalyzed by bacterial RNase P: analysis of precursor tRNAs with a single 3′-S-phosphorothiolate internucleotide linkage. *Nucleic Acids Res.* **28,** 720–727 (2000).
64. Westhof, E. & Altman, S. Three-dimensional working model of M1 RNA, the catalytic RNA subunit of ribonuclease P from *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **91,** 5133–5137 (1994).
65. Harris, M. E., Kazantsev, A. V., Chen, J. L. & Pace, N. R. Analysis of the tertiary structure of the ribonuclease P ribozyme-substrate complex by site-specific photoaffinity crosslinking. *RNA* **3,** 561–576 (1997).
66. Frank, D. N., Adamidi, C., Ehringer, M. A., Pitulle, C. & Pace, N. R. Phylogenetic-comparative analysis of the eukaryal ribonuclease P RNA. *RNA* **6,** 1895–1904 (2000).
67. Li, Y. & Altman, S. A subunit of human nuclear RNase P has ATPase activity. *Proc. Natl Acad. Sci. USA* **98,** 441–444 (2001).
68. Xiao, S., Houser-Scott, F. & Engelke, D. R. Eukaryotic ribonuclease P: increased complexity to cope with the nuclear pre-tRNA pathway. *J. Cell. Physiol.* **187,** 11–20 (2001).
69. Tesmer, J. J. *et al.* Two-metal-ion catalysis in adenylyl cyclase. *Science* **285,** 756–760 (1999).
70. Wyckoff, H. W. *et al.* The three-dimensional structure of ribonuclease-S. Interpretation of an electron density map at a nominal resolution of 2 Å. *J. Biol. Chem.* **245,** 305–328 (1970).
71. Drum, C. L. *et al.* Structural basis for the activation of anthrax adenylyl cyclase exotoxin by calmodulin. *Nature* **415,** 396–402 (2002).
72. Treiber, D. K. & Williamson, J. R. Exposing the kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.* **9,** 339–345 (1999).
73. Thirumalai, D. & Woodson, S. A. Maximizing RNA folding rates: a balancing act. *RNA* **6,** 790–794 (2000).
74. Thirumalai, D., Lee, N., Woodson, S. A. & Klimov, D. Early events in RNA folding. *Annu. Rev. Phys. Chem.* **52,** 751–762 (2001).
75. Treiber, D. K. & Williamson, J. R. Beyond kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.* **11,** 309–314 (2001).
76. Zhuang, X. *et al.* A single-molecule study of RNA catalysis and folding. *Science* **288,** 2048–2051 (2000).
77. Liphardt, J., Onoa, B., Smith, S. B., Tinoco, I. J. & Bustamante, C. Reversible unfolding of single RNA molecules by mechanical force. *Science* **292,** 733–737 (2001).
78. Russell, R. *et al.* Exploring the folding landscape of a structured RNA. *Proc. Natl Acad. Sci. USA* **99,** 155–160 (2002).
79. Caprara, M. G., Mohr, G. & Lambowitz, A. M. A tyrosyl-tRNA synthetase protein induces tertiary folding of the group I intron catalytic core. *J. Mol. Biol.* **257,** 512–531 (1996).
80. Weeks, K. M. & Cech, T. R. Assembly of a ribonucleoprotein catalyst by tertiary structure capture. *Science* **271,** 345–348 (1996).
81. Chanfreau, G. & Jacquier, A. An RNA conformational change between the two chemical steps of group II self-splicing. *EMBO J.* **15,** 3466–3476 (1996).
82. Cohen, S. B. & Cech, T. R. Dynamics of thermal motions within a large catalytic RNA investigated by cross-linking with thiol-disulfide interchange. *J. Am. Chem. Soc.* **119,** 6259–6268 (1997).
83. Profenno, L. A., Kierzek, R., Testa, S. M. & Turner, D. H. Guanosine binds to the *Tetrahymena* ribozyme in more than one step, and its 2′-OH and the nonbridging pro-Sp phosphoryl oxygen at the cleavage site are required for productive docking. *Biochemistry* **36,** 12477–12485 (1997).
84. Murchie, A. I., Thomson, J. B., Walter, F. & Lilley, D. M. Folding of the hairpin ribozyme in its natural conformation achieves close physical proximity of the loops. *Mol. Cell* **1,** 873–881 (1998).
85. Andersen, A. A. & Collins, R. A. Rearrangement of a stable RNA secondary structure during VS ribozyme catalysis. *Mol. Cell* **5,** 469–478 (2000).
86. Noller, H. F., Hoffarth, V. & Zimniak, L. Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* **256,** 1416–1419 (1992).
87. Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289,** 905–920 (2000).
88. Welch, M., Chastang, J. & Yarus, M. An inhibitor of ribosomal peptidyl transferase using transition-state analogy. *Biochemistry* **34,** 385–390 (1995).
89. Nissen, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289,** 920–930 (2000).
90. Polacek, N., Gaynor, M., Yassin, A. & Mankin, A. S. Ribosomal peptidyl transferase can withstand mutations at the putative catalytic nucleotide. *Nature* **411,** 498–501 (2001).
91. Thompson, J. *et al.* Analysis of mutations at residues A2451 and G2447 of 23S rRNA in the peptidyltransferase active site of the 50S ribosomal subunit. *Proc. Natl Acad. Sci. USA* **98,** 9002–9007 (2001).
92. Murray, J. M. & Doudna, J. A. Creative catalysis: pieces of the RNA world jigsaw. *Trends Biochem. Sci.* **26,** 699–701 (2001).
93. Kumar, R. K. & Yarus, M. RNA-catalyzed amino acid activation. *Biochemistry* **40,** 6998–7004 (2001).
94. Illangasekare, M. & Yarus, M. Specific, rapid synthesis of Phe-RNA by RNA. *Proc. Natl Acad. Sci. USA* **96,** 5470–5475 (1999).
95. Illangasekare, M. & Yarus, M. A tiny RNA that catalyzes both aminoacyl-RNA and peptidyl-RNA synthesis. *RNA* **5,** 1482–1489 (1999).
96. Collins, C. A. & Guthrie, C. The question remains: is the spliceosome a ribozyme? *Nature Struct. Biol.* **7,** 850–854 (2000).
97. Yean, S. L., Wuenschell, G., Termini, J. & Lin, R. J. Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* **408,** 881–884 (2000).
98. Valadkhan, S. & Manley, J. L. Splicing-related catalysis by protein-free snRNAs. *Nature* **413,** 701–707 (2001).

Acknowledgements

We acknowledge D. Battle for extensive help with figure preparation, and V. Rath for comments on the manuscript.