Catalytic RNA: Structure and Mechanism

Thomas R. Cech

Howard Hughes Medical Institute, Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215, U.S.A.

In explaining the structure of the parts, if a teacher would be of real sense, he must take care not merely to describe, but to show or demonstrate every part. What the student acquires in this way is solid knowledge arising from the information of his senses.

William Hunter, born 1718

The Hopkins Lecturer is given the charge of explaining how advances in other fields of science have had an impact on his or her own field of biochemistry. A moment's thought revealed that it was not at all difficult to think of examples where other scientific fields have provided necessary groundwork for the study of RNA catalysis or RNA biochemistry in general. Rather, it was difficult to choose among all the possibilities. I have chosen to provide some examples of how the fields of enzymology, synthetic chemistry and even protein crystallography have paved the way for our own research. Of course, there is interchange in both directions, and the Hopkins Lecturer is also charged with explaining how advances in his or her own field have had an impact on biochemistry as a whole. In the case of RNA catalysis, we see spin-offs for research into the origin of life, for the understanding of cellular processes including protein synthesis and mRNA splicing and potentially in the field of medicine.

The system

In *Tetrahymena thermophila*, all of the genes for ribosomal RNA are interrupted by a 413 base-pair intervening sequence (IVS) or intron. This intron is transcribed along with the adjacent exons (mature rRNA sequences) to give a large precursor RNA, which is subsequently spliced. *In vitro* the RNA is self-splicing: the intron excises itself from the precursor and ligates the adjacent exons in the absence of any protein [1, 2]. This intron is a member of a phylogenetically-diverse family called group I [3, 4], which now consists of more than 100 sequenced examples. Many of these have been shown to be self-splicing *in vitro*. In some cases the catalysis provided by the RNA is known to be supplemented *in vivo* by proteins, which are presumed to stabilize the active structure of the intron [5].

The mechanism of splicing (Figure 1) involves two transesterification reactions, each of which switches the substituents that are joined in a phosphodiester bond. Guanosine (or a guanine nucleotide such as GTP) is bound by the intron and

Abbreviations used: E, ribozyme; IVS, intervening sequence; P, product of ribozyme-catalysed cleavage consisting of nucleotides 5' to the cleavage site; S, nucleic acid substrate; sn RNAs, small nuclear RNAs.
Figure 1
Pathway for self-splicing of Tetrahymena pre-rRNA via two transesterification steps

Further details are given in the text. Small letters denote the exons; capital letters, IVS; italic G, exogenous guanosine or GTP. From [42], with permission from Nature; copyright (1990) Macmillan Magazines Ltd.

\[ \text{pre-rRNA} + \text{G}_{\text{OH}} \rightarrow \text{GGGAGG} \]

Step 1

\[ \text{IVS} \quad \text{Step 2} \quad \text{GGGAGG} \quad \text{AG} \]

\[ + \quad \text{cucucucu} \quad 3' \]

ligated exons

acts as a nucleophile, initiating the first transesterification reaction [6, 7]. The first G then leaves the guanosine-binding site and the G at the 3' end of the intron enters. In the second chemical step, the 3' OH group at the end of the first exon attacks the 3' splice site, ligating the exons and releasing the intron.

Shortened forms of the Tetrahymena intron have been engineered to delete the splice sites but to retain the guanosine-binding site, the 5' exon-binding site and other elements of the catalytic apparatus. They cleave or ligate exogenous RNA or DNA substrates with multiple turnover, thereby acting as RNA enzymes [8-10]. One such system (Figure 2) performs an endonuclease reaction analogous to the first step of pre-rRNA self-splicing. Cleavage can occur by transesterification with guanosine serving as the nucleophile or by hydrolysis with water or OH− replacing guanosine. The hydrolysis reaction is chemically similar to that catalysed by another catalytic RNA, the RNA subunit of RNAase P [11].

Enzymology
The artificial enzymic reaction shown in Figure 2 has major advantages for mechanistic studies. As in the self-splicing RNA system, the concentration of the nucleophile (G) can be varied, but unlike the self-splicing system, the concentrations of ribozyme (E) and nucleic acid substrate (S) can be varied independently also. This allows elementary rate constants to be measured. Furthermore, S can be an oligonucleotide and therefore can be synthesized chemically. The ability to synthesize RNA with specific functional group substitutions allows the details of the ribozyme-substrate interactions to be dissected.

We have borrowed concepts and techniques from protein enzymology, including pulse-chase kinetic analysis [12], to determine the individual rate constants and equilibrium constants required to describe the cleavage of RNA by the Tetrahymena ribozyme [13, 14]. The rate-limiting step of the reaction depends on the concentrations of the reacting species. Consider the cleavage of a 'matched' RNA substrate S, that is, one that forms a continuous series of base pairs with the binding site on the ribozyme (Figure 2). Under single-turnover conditions, when S is subsaturating and G is saturating (so-called \((k_{\text{cat}}/K_M)^S\) conditions), the rate-limiting step is the binding of S. As shown in the free energy profile in Figure 3, this binding step is
Pathway for cleavage of an oligonucleotide substrate by the L-21 Scal ribozyme, a shortened version of the Tetrahymena intervening sequence RNA [10]

From [42], with permission from Nature; copyright (1990) Macmillan Magazines Ltd.

Figure 3
Free energy diagram for reaction of unbound RNA substrate (S) and the ribozyme–guanosine complex (E₅) in the case of matched RNA substrate and wild type ribozyme

Binding of S is the rate-limiting step under these subsaturating S or \((K_{\text{cat}}/K_{\text{M}})^s\) conditions. Binding constants and rate constants used to construct this diagram are from [13]. Figure from [16]. Copyright Cell Press.

The highest barrier along the reaction co-ordinate. Thus every substrate molecule that binds goes on to be cleaved in the chemical step. When S is saturating and G is subsaturating \((k_{\text{cat}}/K_{\text{M}})^s\) conditions), the chemical step appears to be partially or fully rate-limiting. Some of the evidence for this proposal has come from the substitution of a phosphorothioate at the cleavage site, which decreases \((k_{\text{cat}}/K_{\text{M}})^s\) by an amount similar to that seen in model chemical reactions [15]. Finally, under multiple-turnover conditions with both S and G saturating (so-called \(k_{\text{cat}}/K_{\text{M}}\)) conditions), the release of P (the 5' terminal fragment of the oligonucleotide) from the active site is rate-limiting.

The rate constant calculated for the chemical step is about 350 min⁻¹, approximately 10¹¹-fold greater than that calculated for the uncatalysed hydrolysis of a phosphodiester bond in RNA. This increase in rate is comparable to that achieved by many protein enzymes.

The availability of the kinetic framework for the endonuclease reaction has been critical for proper interpretation of many observations regarding catalysis by this ribozyme. One example concerns a series of deletion and insertion mutations in a linker region that joins the substrate-binding site of the ribozyme with its catalytic centre [16]. Under multiple-turnover conditions with saturating RNA substrate and a high guanosine concentration, one of these variant ribozymes, named ‘−2A’, showed a dramatic increase in the rate of RNA cleavage and at the same time a 50-fold increase in specificity of cleaving the correct sequence over a sequence that would form a single mismatch when bound to the substrate-binding site of the ribozyme. On the other hand, with subsaturating matched S \((k_{\text{cat}}/K_{\text{M}})^s\) conditions), the cleavage rate of the −2A ribozyme was indistinguishable from that of the wild type. All these superficially incongruous features are explained in that the variant ribozyme binds both the RNA substrate S and the cleavage product P much more weakly than does the wild-type ribo-
zyme. Because the multiple-turnover reaction of both these ribozymes at saturating $S$ is limited by product dissociation, the weaker binding of $P$ results in faster turnover. On the other hand, $(k_{\text{on}}/K_M)^P$ is limited by the initial binding of matched $S$, which is essentially a base-pairing reaction and is not perturbed by the mutation. Finally the increased specificity of the $-2\Delta$ ribozyme is also due to the increased off-rate of $S$. The wild-type ribozyme cleaves essentially every RNA molecule, matched or mismatched, that binds, because these substrates have similar rate constants for association with the ribozyme and because cleavage occurs faster than the substrate can dissociate. In contrast the $-2\Delta$ mutant has a much faster substrate dissociation; the more weakly bound mismatched substrate can now dissociate before it is cleaved, thereby enhancing cleavage specificity.

**Synthetic chemistry**

The availability of automated nucleic-acid synthesis [17, 18] has completely changed the way in which our experiments on ribozyme catalysis are performed. Oligonucleotide substrates of defined length and sequence can be synthesized in large quantity and the effects of single base substitutions can be assessed. While previously many of these substrates could be made by enzymic transcription [9, 10, 19], each RNA polymerase has specific sequence requirements near the start site of transcription that limit the sort of sequences that can be transcribed with high efficiency. Thus chemical synthesis is much more versatile than enzymic synthesis. Furthermore, chemical synthesis allows non-standard bases, sugars and internucleotide bonds to be substituted at a specific position or positions within the chain.

By synthesizing chimeric substrates, in which a single deoxyribose sugar was inserted at successive positions along an otherwise ribose chain, the contribution of individual 2' OH groups both to binding and to reactivity could be assessed. The 2' OH groups at positions $-2$ and $-3$, preceding the cleavage site at position $-1$, were shown to contribute 1-2 kcal/mol (4.2-8.4 kJ/mol) to binding energy [20, 21]. Evidence for a contact between a nucleotide within the catalytic core of the ribozyme (A302) and the OH group at position $-3$ in the substrate was obtained subsequently [22]. The 2' OH group at position $-1$, adjacent to the cleavage site, does not contribute much to binding but makes a very large contribution to transition state stabilization (D. Herschlag, F. Eckstein and T. R. Cech, unpublished work). The rate effect upon substitu-

tion is consistent with the ribose 3' oxyanion being a more stable leaving group than the corresponding deoxyribose oxyanion.

Oligonucleotide synthesis has led also to the identification of an active-site metal ion and to the quantitation of the contribution of this metal ion to catalysis. Cosstick and Vyle [24] recently described the synthesis of DNA in which a bridging-oxygen atom was substituted by sulphur. This is quite different from the standard phosphorothioate, in which a non-bridging-oxygen atom is substituted. A ribozyme substrate containing this novel 3' thio linkage was synthesized and was found to be very poorly cleaved by the *Tetrahymena* ribozyme under the normal Mg$^{2+}$-containing reaction conditions. However replacement of Mg$^{2+}$ with Mn$^{2+}$ resulted in restoration of cleavage activity. Mg$^{2+}$ has a very large preference for co-ordinating to oxygen rather than sulphur, whereas Mn$^{2+}$ has much less of a preference. Thus the ribozyme cleavage results indicated a metal ion interacting with the 3' oxygen or sulphur atom, stabilizing the developing negative charge in the transition state and thereby contributing to catalysis [25] (see Figure 4).

The way in which the ribozyme uses a Mg$^{2+}$ ion for catalysis may be analogous to the role of metal ions in certain protein enzymes, as first proposed by Steitz and coworkers [26]. The exonuclease activity of the Klenow fragment of DNA polymerase, alkaline phosphatase and HIV reverse transcriptase all have two active-site metal ions. One of the metal ions has been proposed to be co-ordinated to the 3' oxygen atom of the leaving group [26], the very interaction observed with the *Tetrahymena* ribozyme. The second metal ion is located below the equatorial plane of the transition state, where it presumably helps activate the nucleophile in a nucleic acid cleavage reaction. The possibility of a second metal ion in an analogous location in the *Tetrahymena* ribozyme is under investigation. Thus the field of protein crystallography is having an impact on our thinking about RNA catalysis. In the opposite direction the studies of RNA catalysis may have an impact on the understanding of the protein enzymes with two active-site metal ions, because some of the mechanistic information gleaned from the ribozyme reaction is not yet known for these protein enzymes. For example in the ribozyme reaction it is known that the metal ion associated with the 3' oxygen atom of the leaving group has very little effect on substrate binding, but specifically stabilizes the transition state. Furthermore the amount of catalysis attributable to the metal ion has been calculated; the single metal ion...
Figure 4
Ribozyme-substrate interactions in the transition state for the endonuclease reaction catalysed by the Tetrahymena ribozyme

Shading, 3-dimensional surface of the ribozyme; dashed lines, hydrogen bonds or metal-ion co-ordination; dotted P-O bonds, bonds partially formed or partially broken in the transition state. From [43].

accounts for $\sim 10^7$ of the $\sim 10^{11}$ fold rate enhancement in the Tetrahymena ribozyme [25].

Repaying the debt
The research on RNA catalysis has benefited repeatedly from advances in other fields, some of which were described above. It has been satisfying to see that the studies on RNA catalysis have also had some impact elsewhere. In the origin of life field, a long-standing controversy regards which came first, the protein or the nucleic acid, the function or the information. We now see that RNA can catalyse biochemical reactions and, as has been known for a much longer time, can serve as a reputable informational molecule. This dual nature of RNA makes it an attractive candidate for a primordial self-replicating system [27–30].

Protein synthesis takes place on ribosomes, which contain both protein and nucleic acid components. Evidence has been accumulating that the RNA component is more than just a scaffold to organize critical proteins, and may participate directly in catalysing protein synthesis. For example, drug-resistant mutations of the ribosome have been mapped to ribosomal RNA, and footprinting studies have shown that elongation factors, transfer RNAs and drug molecules that influence protein synthesis lead to the protection from chemical modification of specific nucleic acid bases [31]. With the discovery of catalytic RNA, it seemed possible that ribosomal RNA might be the catalyst for peptidyl transfer. Consistent with such a view, extensively deproteinized ribosomes from Thermus aquaticus remain capable of carrying out the transfer of an amino acid from the 3' end of a transfer RNA fragment to puromycin [32]. Also consistent with a broader role for RNA in catalysis is the observation that the Tetrahymena ribozyme has modest activity as an aminoacyl esterase [33].

Nuclear mRNA splicing takes place on the spliceosome, a complex assembly of small nuclear ribonucleoprotein particles (snRNPs) and other proteins assembled on the pre-mRNA. Might the snRNAs provide the active site for the two transesterification reactions required for mRNA splicing? Circumstantial evidence for this proposal is accumulating, although no rigorous demonstration of RNA catalysis by snRNAs is yet in hand [34, 35].

In a more general sense, the studies on catalytic RNA have had an impact on the general perception of RNA structure. The first idea about RNA structure, derived from thinking about the coding capacity of messenger RNAs, was as a linear sequence of nucleotides. As it became apparent that many RNAs relied on a specific structure for their activity, this spaghetti model of RNA structure was replaced by secondary-structure models in which the RNA was shown as a series of stems and loops with connecting single-stranded regions. But RNA catalysis clearly requires a specific three-dimensional structure, because RNA elements distant in the secondary structure must come into proximity for the reaction to occur. In the case of the group I introns, the outlines of this structure are rapidly being clarified [36, 37]. Generalizing, many RNA molecules may have specific, compact tertiary structures. Thus at low resolution many RNA molecules may be envisioned as resembling globular proteins, although of course the types of interactions that hold together RNA and protein are very different in detail [38].
Finally, there is hope that studies on RNA catalysis may have an impact on medicine. A number of plant infectious agents, including certain viroids, virusoids and satellite RNAs, contain a nucleotide sequence that folds to form a 'hammerhead' structure that undergoes self-cleavage [39]. Small RNAs containing this catalytic centre bind to and cleave exogenous RNAs [40, 41]. Such hammerhead ribozymes have been engineered to recognize mRNAs and viral RNAs and have been shown to cleave them under controlled conditions in vitro and even in vivo. There is now considerable activity in the area of delivering small ribozymes both to cells and to animals, with the aim of cleaving and thereby inactivating RNA molecules involved in human disease.

I am grateful to a large number of skilled coworkers who have done the experiments and contributed to the ideas of RNA catalysis. In particular, I thank Dan Herschlag, Joe Piccirilli and Anna Marie Pyle for developing some of the ideas presented here. I also thank Pascale Legault and Tim McConnell for comments on this manuscript and Alice Sirimarco for help in its preparation. I am an Investigator at the Howard Hughes Medical Institute and a Professor of the American Cancer Society. Our research on RNA catalysis is also supported by a grant from the National Institutes of Health.

23. Reference deleted

Received 11 November 1992