The origins of RNA catalysis in ribozymes

David M.J. Lilley

Cancer Research UK Nucleic Acid Structure Research Group, School of Life Sciences, MSI/WTB Complex, The University of Dundee, Dundee, U.K., DD1 5EH

The discovery of RNA catalysis provided a paradigm shift in biology, insight into the evolution of life on the planet and a challenge to understand its mechanistic origins. RNA has limited catalytic resources that must be used to maximal effect. Consequently, RNA catalysis tends to be multifactorial, with several processes contributing to an overall significant enhancement of reaction rate. These include general acid–base catalysis, electrostatic effects, and substrate orientation and proximity. The main players are the RNA nucleobases and bound metal ions. Although most ribozymes carry out phosphoryl transfer, the same considerations appear to apply to peptidyl transfer in the ribosome.

Until 20 years ago it was firmly believed that proteins were the only catalytic macromolecules in biology. However, this conception was overturned with the discovery of the first catalytic RNA molecules, or ribozymes [1,2]. Explaining how molecules as chemically dull (compared with proteins anyway) as RNA could possibly function as pseudo-enzymes proved a great challenge to the biological chemist. In the intervening two decades we have learned a great deal about the origins of catalysis in RNA, yet most aspects have been controversial and only now is some semblance of a general picture emerging.

The occurrence of RNA catalysis

In the majority of cases the substrates of ribozymes are RNA, and most ribozymes serve an RNA-processing function. The simplest are the nucleolytic ribozymes, which are required for the site-specific cleavage and ligation of replication intermediates [3,4]. Although the chemical reaction is the same in each case, there are four distinct classes of nucleolytic ribozymes that have unrelated structures: hammerhead and hairpin ribozymes, mostly found in plant viruses, the Varkud satellite (VS) ribozyme in fungal mitochondria and the hepatitis delta virus (HDV) ribozyme, which is present in a human pathogen. Many ribozymes are involved in the self-splicing of introns in a wide range of species, and can be divided into two groups (group I and group II) [1,5].

The nucleolytic ribozymes and the group I and II ribozymes all perform different kinds of phosphoryl-transfer reactions, in which a transesterification reaction results in breakage of the backbone in the first step (Fig. 1). This is brought about by the attack of a ribose hydroxyl oxygen atom on the phosphodiester bond, but is different for each type of ribozyme. The nucleophile is the adjacent 2'-hydroxyl in the nucleolytic ribozymes, whereas it is a remote 2'-hydroxyl in group II and the 3'-hydroxyl of exogenous guanosine in group I introns. For both group I and group II ribozymes there is a subsequent reaction that excises the intron and joins the exons precisely. The

Fig. 1. Phosphoryl-transfer reactions in ribozymes. (a) In the nucleolytic ribozymes the phosphorus is attacked by the adjacent 2'-hydroxyl with departure of the 5'-oxygen, generating a cyclic 2',3' phosphate. (b) In the first step of the group II intron reaction, the scissile phosphorus is attacked by the 2'-hydroxyl of an adenosine that is located in the middle of the intron to be excised. (c) By contrast, in the group I intron, the attack comes from the 3'-hydroxyl of an exogenous guanosine that is bound non-covalently by the ribozyme. Introns are shown in red, exons in blue and exogenous guanosine in green.
similarity of the chemistry of the group II ribozyme to mRNA splicing has suggested to many that the spliceome will turn out to be a ribozyme. There are some experimental indicators of this [6–9], but no smoking gun as yet, and at present we cannot exclude a role for protein (such as a histidine side chain perhaps) participating in the catalysis. Ribonuclease P (RNaseP) is a nucleoprotein ribozyme that uniquely carries out a hydrolytic cleavage reaction to remove the 5′ end of pre-mRNA [2]. The exception to the rule that ribozymes generally act upon RNA is the remarkable observation that the peptidyltransferase reaction of the ribosome is catalyzed by the 23S rRNA component of its 50S subunit [10,11].

**Why is RNA catalysis important?**

There are several reasons why ribozymes are worth studying. First, RNA catalysis is a genuine phenomenon in biology and, therefore, a bona fide subject of interest to the biologist. For quite a time it appeared that RNA catalysis was limited to some rather special cases of a few plant pathogens and some splicing reactions occurring in slimy things found in ponds. But RNaseP is a ubiquitous processing activity, and it might turn out that mRNA splicing is performed by a ribozyme. The demonstration that the ribosome is a ribozyme shows us that what is arguably the most important reaction of the cell – the condensation of amino acids into polypeptides – is catalyzed by RNA. This really elevates RNA catalysis to mainstream cell function.

Second, ribozymes might have played a key role in the development of life on this planet. There is clearly a fundamental ‘chicken and egg’ problem in imagining how life came into being in anything like its present form. Even the simplest translation system would surely be too complex to exist at an early stage. This problem can be circumvented if primitive life passed through a stage at which RNA was both the informational and the catalytic macromolecule [12,13]. Those of us studying ribozymes might like to think that we are peering back 3.6 billion years into this ‘RNA world’ in our more fanciful moments! The ribosome (perhaps RNaseP too) probably represents a molecular fossil from this period that still exists inside all cells, although because these ribozymes are still functional today, a kind of molecular coelacanth might be a better analogy.

Third, perhaps ribozymes can tell us something more general about biocatalysis. Compared with protein, RNA is a rather poor prospect as a catalyst, yet somehow it achieves some respectable rate accelerations. Perhaps we can draw some general mechanistic conclusions from Nature’s ‘stripped-down’ catalysts.

**RNA folding and catalysis**

Just as a protein must fold into its active conformation, a ribozyme must fold into the conformation that provides the environment in which catalysis can proceed [14–16]. Moreover, the local RNA structure might be a direct participant in the catalytic mechanism. RNA is a polyelectrolyte and, therefore, counterions are required to neutralize charge, and divalent metal ions are key players in mediating folding into the active structure. Two kinds of ion binding can be distinguished – diffuse binding and site-specific binding. In diffuse binding, metal ions are bound by electrostatic forces in regions of high potential while retaining their inner-solvent coordination sphere. Although high occupancy is possible, such ions will exhibit rapid exchange. If one or more of the inner-sphere water molecules are substituted for RNA ligands (e.g. a phosphate oxygen atom), then site-specific binding results, with slower exchange and a greater selectivity for ion size and charge.

**Origins of catalysis in the nucleolytic ribozymes**

So how does RNA act like an enzyme? Let us concentrate on the nucleolytic ribozymes; there are crystal structures for three of the four classes, and some of the most detailed mechanistic studies have been carried out on these species. They are, therefore, very good subjects for trying to deduce the basic mechanisms of catalysis.

The reaction mechanism for these ribozymes is an $S_{N}2$ reaction resulting from nucleophilic attack by the 2′-O on the 3′-P, with departure of the 5′-O (Fig. 2). The product is cyclic 2′,3′-P [17], and the reaction is accompanied by an inversion of configuration at the phosphate (revealed if one of the non-bridging oxygen atoms is replaced by sulfur [18,19]). The transition state is an oxyprophosphorane intermediate with a trigonal bipyramidal structure.

![Fig. 2](http://tibs.trends.com)
containing the attacking 2'-O and departing 5'-O atoms at apical positions. Strictly speaking, the symmetrical oxyporphosphate might be a high-energy intermediate flanked by associative and dissociative transition states of higher, and different, energies, but it is probably close enough that this distinction need not concern us here. The reactions carried out by the group I and II introns are basically similar, except that the attacking hydroxyl is different from the adjacent 2'-O and, thus, similar considerations will apply to their catalysis.

The same reaction can be brought about by specific base catalysis. Indeed, the reason why RNA is relatively unstable compared with DNA is that the 2'-hydroxyl is a kind of ‘dagger poised at its heart’. The rates of the base-catalyzed reaction proceeding in simple unstructured RNA species (e.g. dinucleotides) have been measured at close to neutral pH [20]. Comparison with rates measured for the ribozymes (typically ~0.1–1 min⁻¹) show that the ribozymes achieve a rate enhancement of ~10⁵–10⁶. It is this catalytic power that requires explanation.

Does the mechanism of the ribozyme reaction suggest how catalysis might be achieved? How could this reaction be accelerated by an RNA catalyst – can we identify points that might be amenable to catalytic intervention? Three broad strategies are possible:

1) General acid–base catalysis. A hydroxyl group is not a strong nucleophile, and removal of the proton by a general (Brønsted) base would create a stronger alkoxide-ion nucleophile. The departure of the oxanion would be facilitated by protonation, requiring a general acid. Of course, in the reverse (ligation) reaction the acid and base would exchange roles by the principle of microscopic reversibility, whereby the elementary steps in enzyme-catalyzed reactions are mechanistically reversible. A clear precedent for these processes is set by the enzyme RNaseA, which performs exactly the same transesterification backbone cleavage, catalyzed by the imidazole side chains of two histidine residues. One (His12) acts as the general base to deprotonate the 2'-hydroxyl, while the other (His119) protonates the departing 5'-oxyanion as a general acid.

2) Charge stabilization. The transition state of the reaction is doubly charged and could be stabilized electrostatically by juxtaposition of positive charge.

3) Conformational effects – proximity and orientation. Passage into a trigonal bipyramidal transition state requires an in-line attack, aligning the 2'-O, 3'-P and 5'-O. However, the usual conformation of RNA could hardly be more different from this. The local environment of the ribozyme could induce conformational changes to facilitate the trajectory into the in-line transition state. More generally, positioning and orientation of substrates is a significant factor in achieving rate enhancements in enzymes, and this is a strategy that could be employed in ribozymes.

What resources are available to RNA for catalysis?
Proteins are Nature’s chemistry set, with side chains that cover a wide range of aqueous chemistry. By comparison, RNA has limited resources, with four similar heterocyclic nucleobases and a relative paucity of functional groups. So what are the options available?

Metal ions
Without question, metal ions play a crucial role in the folding of RNA molecules, thus creating the active geometry of ribozymes. And because the local conformation might be an important factor in catalysis, this could be regarded as a secondary effect of the metal ions. But, for a long time it was thought that bound metal ions would play a direct role in the chemistry of ribozyme reactions and, at one stage, ribozymes were thought to be special examples of metalloenzymes.

Metal ions could play several different direct roles in catalysis. First, metal-coordinated water could act in general acid–base catalysis, as originally deduced for the cleavage of tRNA in lead ions [21]. The action of solvated metal hydroxide was deduced for the hammerhead ribozyme from the linear pH dependence of the rate [22]. Second, a metal ion could form an inner-sphere complex with the RNA, acting as a Lewis acid to assist deprotonation of the attacking nucleophile or stabilization of the charged leaving group. Third, a metal ion could stabilize the charged transition state electrostatically, possibly by binding to one of the non-bridging oxygen atoms. Several protein enzymes use a two-metal-ion strategy that involves binding one metal ion to the attacking nucleophile while the other binds to the leaving group [23].

Two, in particular, interact with the attacking and leaving oxygen atoms.

By contrast, evidence for a direct role of metal ions in the reactions of the nucleolytic ribozymes has grown weaker in recent years. Experiments on the hairpin ribozyme have shown that either of the non-bridging oxygen atoms of the scissile phosphate can be replaced by sulfur with no impairment of cleavage rate [25–27]. This is in contrast to the hammerhead ribozyme, which exhibits a large thio effect on the Rₚ non-bridging oxygen atom [28]. It was also shown that the hairpin ribozyme retains considerable activity when magnesium ions are replaced by substitutionally inert hexammine cobalt (III) ions [25–27]. This excludes a role for metal ions in general acid–base catalysis, as well as action as a site-bound Lewis acid. Fourth, it was shown that respectable levels of activity could be attained in the hammerhead, hairpin and VS ribozymes in the presence of high concentrations of monovalent ions such as lithium and ammonium [29], which again seems to exclude a requirement for site-specific binding of the ion.

Nucleobases
In principle, nucleobases might participate directly in the catalytic chemistry. For example, the ring nitrogens N1 of adenine and N3 of cytosine could become protonated. The
nucleobases might, therefore, participate in general acid–base chemistry, or protonated bases might stabilize negatively charged transition states electrostatically. However, the pKₐ values of these bases are far from neutrality (3.5 and 4.2 for adenine and cytosine, respectively), whereas it is important that the pKₐ is close to 7.0 for a catalyst operating near neutral pH. Because the conjugate base of a strong acid is a weak base it will act poorly as a general base, and only one cytosine molecule in 10⁴ will be in the acidic protonated form. So, to be useful catalytically, the pKₐ must be perturbed by the local environment to be closer to 7.0. This is entirely possible

Fig. 3. The hairpin ribozyme. (a) The secondary structure of the ribozyme. The core of the ribozyme comprises two formally unpaired loops carried on adjacent arms of a four-way helical junction. The junction is dispensable for activity, but folding is orders of magnitude more efficient with the junction present. Ribozyme cleavage occurs at the position indicated by the red arrow. (b) The structure in the crystal, determined by Ferré d’Amare and coworkers [32], coloured in the same manner as (a). The interface for the interaction between the loops is complex, but includes the extrusion of G₈₊₁, and insertion into the loop of arm B, where it forms a conventional basepair with C₂₅. (c) The local structure at the cleavage site. The extrusion of G₈₊₁ leads to a local backbone conformation that creates an approximately in-line geometry for the attacking 2'‐hydroxyl of A−1. The nucleobase of G₈ is immediately adjacent, and could play a direct role in catalysis. A recent structure from Ferré d’Amare and coworkers [46] containing vanadate as a transition-state analogue shows that the transition state is stabilized by hydrogen bonding additional to that of the ground state.

Fig. 4. The hepatitis delta virus ribozyme. (a) The secondary structure of the product form of the ribozyme; cleavage has occurred at the position shown by the red arrow, 5′ to G₁. The global structure of the ribozyme is a complex, double pseudo-knot. (b) The structure in the crystal, determined by Doudna and coworkers [34], coloured in the same manner as (a). A central deep cleft is formed, enclosed by helices P₁ and P₁.1, the pseudo-knot connections and the lower helix P₄. The cleavage site at G₁ (bold cyan) is found within this cleft, adjacent to the nucleobase of C₇₅ (bold pink). (c) Shows the juxtaposition of G₁ and C₇₅. N₃ of C₇₅ is directed towards the 5′-oxygen of G₁ and is, therefore, well placed to act as the general acid in the cleavage reaction.
in the highly charged context of RNA, and values of 6.0 or higher have been measured by NMR [30].

**Nucleophile alignment in the hairpin ribozyme**

The hairpin ribozyme comprises two formally unpaired loops presented on adjacent arms of a four-way junction (Fig. 3). In the presence of magnesium ions the ribozyme folds, by an intimate association between the loops, to form the active site [31]. This structure has been solved by crystallography, which has revealed the details of a loop–loop interaction [32] (Fig. 4). The base G+1 (immediately 3’ to the scissile phosphate) is extruded from loop A and inserted into a pocket in loop B, where it base-pairs with C25, as predicted earlier [33]. This results in a change in the backbone conformation such that the attacking 2’-hydroxyl of A−1 is now in an approximately in-line geometry. Nucleobase catalysis might also be important. G8 has been suspected to play such a role, and is found adjacent to the scissile phosphate, hydrogen bonded to the phosphate and the nucleophilic hydroxyl.

**Nucleobase catalysis in the HDV ribozyme**

The HDV ribozyme provides the best-characterized example of nucleobase catalysis. The structure of the product has been solved by crystallography [34], revealing a deep cleft containing the site of cleavage (Fig. 4). It also contains an essential cytosine base (C75 in the genomic ribozyme) that is adjacent to the 5’-O of the product. The HDV ribozyme in which C75 was replaced with uracil (C75U) was inactive, but activity could be substantially restored by addition of a high concentration of the base imidazole [35], providing strong evidence for general acid–base catalysis by C75. Careful analysis of the pH dependence of the reaction indicated that the cytosine base was probably acting not as a general base, but as a general acid [36] with a pK_a of ~6.0. General base catalysis was provided by metal-bound water. Perhaps surprisingly, direct measurement of the cytosine pK_a by NMR indicated that this was not significantly elevated [37], raising the possibility that this might only occur as the reaction progresses towards the transition state. Evidence for nucleobase catalysis is also growing rapidly for the VS ribozyme, where A756 is looking like a strong candidate [38,39].

**Peptidyl transferase activity in the ribosome**

The chemistry of the condensation of aminoacyl-tRNA species in the ribosome is a very different from the phosphoryl-transfer reactions. The attacking nucleophile is the α-amino of an amino acid, which attacks an sp^2-hybridized carbon atom of the carbonyl group to yield a tetrahedral intermediate species (Fig. 5); this must be resolved in a subsequent step with the departure of the deacylated tRNA.

Early experiments suggested that the ‘enzyme’ of peptidyl transferase was RNA, not protein [10], and when the structure of the large ribosomal subunit was solved by crystallography this was confirmed. A co-crystallized tetrahedral intermediate analogue is surrounded by the 23S rRNA and the nearest protein is 18 Å from the site [11]. The nucleotide closest to the inhibitor is A2451, which is within hydrogen-bonding distance (Fig. 5), and it has been suggested that this could act as a general base to deprotonate the attacking amine. This was almost immediately contradicted by data from mutants in vivo and in vitro, and a confusing situation resulted for a while.

Recently, a detailed analysis of the pH dependence of the peptidyl bond-forming reaction under conditions where the chemistry was rate limiting has been carried out [40], and shows that an RNA nucleobase with a pK_a of 7.5 contributes to catalysis. Mutation of the suspected crucial adenine (A2451U) leads to a rate reduction and loss of pH dependence, and the residual catalysis is probably due to substrate positioning.

**RNA catalysis – making a lot out of a little**

As a catalytic macromolecule, RNA has rather meagre resources yet manages to make the most of what it has. Certain common catalytic strategies are found, such as the roles of nucleobases and metal ions in general acid–base and electrophilic catalysis, and the importance of
substrate orientation and proximity. Even peptidyl transferase, with a very different underlying chemistry from the phosphoryl-transfer reactions, appears to use some of the same strategies. Few of the individual catalytic strategies by themselves are likely to lead to large rate accelerations. Orientation, for example, can probably not contribute more than a factor of (at most) 100 to rate enhancement, and a similar acceleration is probably available from general acid–base catalysis. Therefore, RNA catalysis is multifactorial, combining different strategies to give the observed overall rate enhancements. Different ribozymes seem to combine the possible factors in different ways, yet often resulting in similar final rates.

Limits to RNA catalysis and the feasibility of an RNA world

Peptidyl transferase notwithstanding, the range of chemistry displayed by contemporary ribozymes is limited. A sustainable RNA world would have required a much wider range of reactions to generate primitive metabolism and to transmit genetic information. One way of examining the possibilities is to use selection methods to try to recreate RNA enzymes with some of those activities. A significant number of ribozymes required for nucleotide synthesis have been generated, such as a phosphoribosyl transferase [41]. Seemingly, the most fundamental requirement would be a template-dependent RNA polymerase and a ribozyme that can transcribe a 15-nucleotide RNA has been obtained [42]. More generally, ribozymes have been selected that catalyze a significant range of relatively exotic chemical reactions, including S-alkylation [43], Michael addition [44] and Diels–Alder electrocyclic addition [45]. Because chemists are relatively new to this game, and Nature has had millions of years to perfect its art, this suggests that the RNA world is probably chemically feasible. Ultimately, the finest achievement of the RNA world was probably the creation of proteins. These then took over most of the catalytic functions, leaving the ribosome as the most permanent monument to a heroic era.

Acknowledgements

I thank Fritz Eckstein for discussion, and Cancer Research-UK and the BBSRC for financial support of work on ribozymes in my laboratory.

References

7 Gordon, P.M. et al. (2000) Metal ion catalysis during the exon-ligation step of nuclear pre-mRNA splicing: extending the parallels between the spliceosome and group II introns. RNA 6, 199–205
14 Bassi, G.S. et al. (1997) Ion-induced folding of the hammerhead ribozyme: a fluorescence resonance energy transfer study. EMBO J. 16, 7481–7489
27 Young, K.J. et al. (1997) Metal ions play a passive role in the hairpin ribozyme catalysed reaction. Nucleic Acids Res. 25, 3760–3766
33 Pinard, R. et al. (1999) Structural basis for the guanosine requirement of the hairpin ribozyme. Biochemistry 38, 16035–16039

---

**Endeavour**

the quarterly magazine for the history and philosophy of science

You can access *Endeavour* online either through your BioMedNet Reviews subscription or via [ScienceDirect](http://reviews.bmn.com), where you’ll find a collection of beautifully illustrated articles on the history of science, book reviews and editorial comment.

featuring

The pathway to the cell and its organelles: one hundred years of the Golgi apparatus
by M. Bentivoglio and P. Mazzarello

Joseph Fourier, the ‘greenhouse effect’ and the quest for a universal theory of terrestrial temperatures
by J.R. Fleming

The hunt for red elixir: an early collaboration between fellows of the Royal Society
by D.R. Dickson

Art as science: scientific illustration 1490–1670 in drawing, woodcut and copper plate
by C.M. Pyle

The history of reductionism versus holistic approaches to scientific research
by H. Andersen

Reading and writing the Book of Nature: Jan Swammerdam (1637–1680)
by M. Cobb

Coming to terms with ambiguity in science: wave–particle duality
by B.K. Stepansky

The role of museums in history of science, technology and medicine
by L. Taub

The ‘internal clocks’ of circadian and interval timing
by S. Hinton and W.H. Meck

The troubled past and uncertain future of group selectionism
by T. Shanahan

A botanist for a continent: Ferdinand Von Mueller (1825–1896)
by R.W. Home

Rudolf Virchow and the scientific approach to medicine
by L. Benaroyo

Darwinism and atheism: different sides of the same coin?
by M. Ruse

Alfred Russel Wallace and the flat earth controversy
by C. Garwood

John Dalton: the world’s first stereochemist
by N.G. Coley

Forensic chemistry in 19th-century Britain
by N.G. Coley

Owen and Huxley: unfinished business
by C.U.M. Smith

Characteristics of scientific revolutions
by H. Andersen

and much, much more . . .


Log on to [http://reviews.bmn.com](http://reviews.bmn.com), hit the ‘Browse Journals’ tab and scroll down to *Endeavour*.