# Ribozymes: aiming at RNA replication and protein synthesis

Alicia J Hager, Jack D Pollard, Jr and Jack W Szostak

The RNA world hypothesis is founded on the idea of an RNA replicase, or self-replicating RNA molecule, and presupposes the later emergence of ribozymes capable of catalyzing the synthesis of peptides. The recent demonstrations of ribozyme-catalyzed template-directed primer extension, and of ribozyme-catalyzed amide bond synthesis, confirm the plausibility of the RNA world, and highlight the steps that remain to be demonstrated in the laboratory.

Address: Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA.

**Chemistry & Biology** September 1996, **3**:717–725

 ISSN 1074-5521

## Introduction

The notion of an 'RNA world' lies at the heart of most modern theories of the origin of life [1]. It is proposed that in the ancient RNA world, ribozymes, rather than protein enzymes, catalyzed the reactions responsible for the maintenance and propagation of life. The plausibility of proposed RNA-world scenarios relies heavily upon the hypothetical existence of two types of ribozymes. First among these is the RNA replicase, an RNA molecule or complex capable of self-replication. It is reasonable to believe that RNA could replicate itself, as it can both store genetic information and catalyze chemical reactions, thereby side-stepping the dilemma of how both catalytic function (currently the domain of proteins) and transmission of information (currently the domain of DNA) could have arisen simultaneously in the common ancestor of all living organisms. At some later point, ribozymes must have evolved to facilitate the transition from an ancient biology based upon nucleic-acid catalysis to modern biology based on protein enzymes. At the least, this would have required a ribozyme capable of catalyzing peptide synthesis from activated amino acids.

Support for RNA-world theories can be found in modern-day cellular mechanisms that involve RNA. For example, self-splicing introns catalyze a chemical transformation that is not very different from that involved in nucleic acid replication; in addition, modern ribosomal RNA is generally thought to be the catalytic core of the ribosome, and may be the evolutionary descendant of a very early peptide-synthesizing ribozyme.

In this article, we review efforts to support the validity of the RNA-world hypothesis by establishing that the key (but so far hypothetical) ribozymes of the RNA world could have actually existed. This approach involves the use of *in vitro* selection, a way of screening very large numbers of nucleic acid sequences for rare functional molecules [2–4].

## RNA replicases

### Conditions for replicase emergence

The initial development of a replicase ribozyme would have required that polynucleotides of a reasonable length (30–70 nucleotides (nt), *vide infra*) be available in the prebiotic world. This is a formidable requirement, so formidable indeed that many prominent researchers in the field ridicule RNA as a plausible prebiotic molecule [5,6]. Considerable attention has therefore been devoted recently to the search for macromolecular structures that retain the desirable properties of RNA but would have been simpler

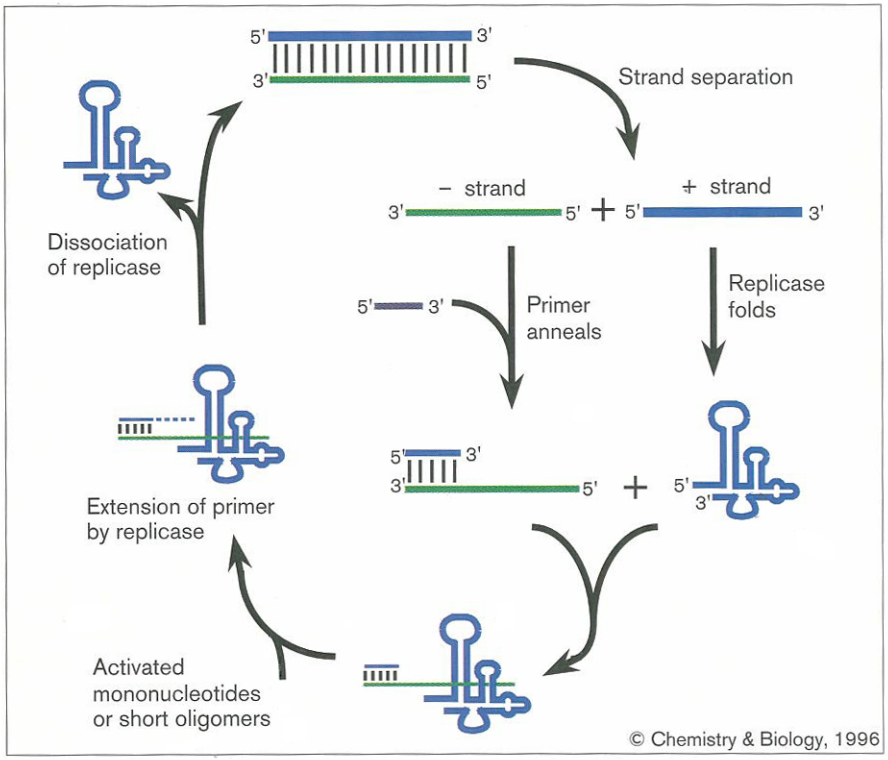
to synthesize under prebiotic conditions, for example, peptide-nucleic acid (PNA) and pyranosyl-RNA (P-RNA). While this is a perfectly reasonable approach, our ignorance of prebiotic chemistry is severe, and it may be premature to dismiss prebiotic RNA entirely. Indeed, some exciting progress has recently been made in the study of the non-enzymatic synthesis and copying of RNA. A chronic problem in the field of prebiotic RNA synthesis has been the short size (10 nucleotides or less) of the oligonucleotides formed by the spontaneous assembly of activated nucleotides in aqueous media. By binding the growing polynucleotide chain to a montmorillonite surface, and repeatedly flushing this surface with fresh solutions of activated monomers, polynucleotides up to 55 nucleotides long have been produced [7]. These populations undoubtedly contain ribozyme catalysts. One major concern with mononucleotide condensation reactions, including the montmorillonite reaction, is the presence of a mixture of 2'-5' and 3'-5' linkages in the product RNA. This heterogeneity is an impediment to subsequent templating activity, but Ertem and Ferris have now demonstrated that even heterogeneously linked oligocytidylates can template the condensation of imidazole-activated monomers to form oligoguanylates in solution [8]. The greater chemical stability of duplex RNA, coupled with the greater thermodynamic stability of all 3'-5' linked duplexes, suggests that 3'-5' linked duplex RNA might accumulate with time in the right environment.

### Self-replication by RNA

To begin RNA-catalyzed RNA replication, either two copies of the replicase or, more likely, one copy of the replicase and one copy of its complementary strand would have been required: one copy to act as a polymerase, and its complement (or a second replicase copy) to act as a template (Fig. 1). The great virtue of the accumulation of double-stranded, as opposed to single-stranded, RNA is therefore obvious — upon strand separation, one copy of a potential replicase could fold up and act as a polymerase, ready to copy its nearby complementary strand, thus initiating the auto-catalytic cycle of replication.

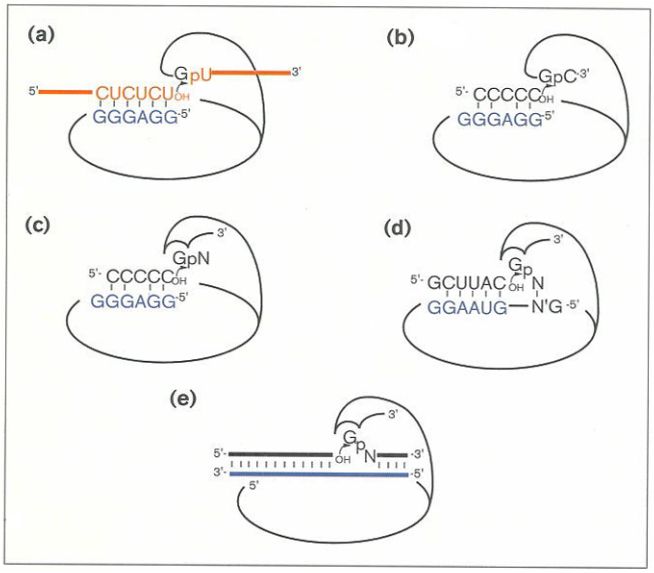
As simple as it seems, a deeper consideration of this scenario immediately engenders further complications. What force could separate the two strands? Long RNA duplexes are notoriously difficult to separate by thermal denaturation. Perhaps a mineral surface with higher affinity for single stranded RNA than duplex RNA could help (L. Orgel, unpublished results), but such effects remain to be demonstrated. And once the strands are separated, some sort of compartmentalization would be needed to prevent the separate RNA strands from drifting apart too quickly. Again, adsorption on a mineral surface could satisfy this requirement, but encapsulation in a membrane vesicle would be more satisfying from a biological perspective. Finally, a successful replicase would have to have a fidelity good enough to copy its own sequence, that is, an

**Figure 1**




Scheme for RNA self-replication. The replicase or '+ strand' (shown in blue) binds to a copy of its complement or '– strand' (shown in green) annealed to a short primer oligomer. The complementary strand serves as template for the primer extension reaction catalyzed by the replicase. The replicase dissociates upon completion of the extension reaction. Following disruption of the newly formed duplex, both the + and – strands are available to serve as templates. Two rounds of copying (+ to –, and – to +) complete a cycle of self-replication.

**Figure 2**




Comparison of reactions catalyzed by the *Tetrahymena* group I intron and derived ribozymes. **(a)** Exon ligation step of the *Tetrahymena* intron self-splicing reaction. The 5' exon and the 3' exon are shown in red. The internal guide sequence (IGS) is shown in blue. **(b)** A single step of the cytidylic-acid disproportionation reaction performed by a ribozyme derived from the *Tetrahymena* group I intron by removal of the exon sequences and some additional nucleotides. **(c)** GpN extension performed by the ribozyme. The guanosine binding site of the intron binds the guanosine of the externally supplied GpN. **(d)** GpN extension with altered IGS sequence and templating of added nucleotide. N' represents the complement of the nucleotide N. **(e)** Ligation of external RNA substrates performed by the ribozyme.

error rate of less than 1/n, where n is the number of important bases in the replicase [9]. The ability of RNA to catalyze complete cycles of auto-catalytic self-replication is far from being established experimentally, but considerable progress has been made in the demonstration of RNA-catalyzed polymerase-like reactions.

### Polymerase-like group I ribozyme reactions

The first attempts to examine the ability of RNA to catalyze RNA polymerization made use of reactions analogous to the exon-ligation step of the self-splicing reaction of group I introns (Fig. 2a). These introns already perform ligations of RNA molecules in the absence of protein, and Zaug and Cech [10] showed that a ribozyme comprising a segment of the *Tetrahymena* group I intron was capable of catalyzing a disproportionation reaction in which cytidylic acid pentamers ($C_5$) were transformed into a heterogeneous population of oligo-C oligonucleotides, with some of the initial $C_5$ strands being extended in a series of transesterification reactions as far as $C_{30}$ (Fig. 2b, see Table 1). Soon thereafter, Been and Cech [11] described an even more interesting polymerization reaction in which guanylyl-(3',5')-nucleotides (GpNs) were used as donors of mononucleotide in the net elongation of $C_5$ to $C_{10}$ (Fig. 2c). The fidelity of the primer extension reaction could not be readily assessed in this system, as the $C_5$ primer could lie down on the template (the internal guide sequence of the ribozyme) in several different positions. By modifying this system so that a mixed-sequence primer would lie down in one defined position on a complementary template, Bartel and Szostak [12] were able to study the extent to which this reaction was template-directed (Fig. 2d). The observed fidelity varied widely with the identity of the template base, and with the concentrations of the competing GpN substrates. On average, and in the presence of saturating concentrations of all four GpNs, the calculated error rate was a disappointing 35 %, far too great to support self-replication.

An alternative to primer extension one nucleotide at a time is the template-directed ligation of oligonucleotide substrates (Fig. 2e). The advantages of this approach are that fewer catalytic cycles are required to generate a long product, and that lower substrate concentrations can be used, because of the tighter binding of the substrate to the template. Although the system becomes artificial when only the correct oligonucleotide substrates are supplied, this approach has provided a useful experimental opening to some of the problems of self-replication. Doudna and Szostak [13] showed that the *Tetrahymena* ribozyme was able to use a wide variety of sequences (including parts of its own sequence) as a template in oligonucleotide assembly reactions (Table 1). Product strands of 40–50 nts could be assembled in four or five steps. To address the question of how the same sequence could function efficiently as both a folded ribozyme and an unfolded template, Doudna *et al.* [14] divided a small derivative of the *sunY* ribozyme into three segments of 59, 75 and 43 nucleotides. Not only could the three short ribozyme subunits anneal to each other and self-assemble into an active ribozyme, but the assembled ribozyme could use additional copies of its own oligonucleotide subunits as templates. Oligonucleotides 8–11 nucleotides in length were ligated together by the ribozyme to form full-length

**Table 1**

**Group I ribozyme reactions related to polymerization.**

| Reaction type | Reaction | Ref. |
|---|---|---|
| Cytidylic acid disproportionation | $C_5 + C_5 + C_5 \rightarrow C_6 + C_4 + C_5$<br>$C_6 + C_4 + C_5 \rightarrow C_7 + 2\,C_4$ | 10 |
| GpN extension | $C_5 + 5\,GpC \rightarrow C_{10} + 5\,G$ | 11 |
| Complementary strand synthesis | G G G → + 3 G | 13 |
| Triplet extension | $N_6 + (2\,AP^a)pNNN \rightarrow N_9 + 2\,AP$ | 15 |
| Exon polymerization | 2 5' exon–3' exon<br>→ 5' exon + 5' exon–3' exon–3' exon | 16 |

[a]2-aminopurine riboside

strands complementary to the ribozyme itself [14]. Doudna and Szostak [15] also found that oligonucleotide substrates as short as tetramers could be used by *sunY*-derived ribozymes in reactions in which a primer was extended three nucleotides at a time. Other group I ribozymes are also capable of extensive oligonucleotide ligation reactions. Burke and coworkers [16] have demonstrated an impressive reaction in which the *Azoarcus* intron uses a 20 nucleotide exon-exon analog as a substrate and generates chains of ligated exons up to 160 nucleotides long (Table 1).

### *In vitro* evolution of group I ribozymes

The poor yields of full-length product in the above experiments suggested that a ribozyme that was both smaller and more active would be a better starting point for development into a self-replicating system. Green and Szostak [17] therefore applied the recently developed *in vitro* selection technique to the optimization of the ligation activity of a short (~140 nucleotides) deletion derivative of the *sunY* ribozyme. This deletion derivative was a superior template because of its less extensive secondary structure. A pool of approximately $2 \times 10^{13}$ sequence variants (mutagenized 5 % at each site) was subjected to rounds of increasingly stringent selection for ligation activity, followed by amplification. A quintuple mutant ribozyme with superior ligation activity was obtained, which also retained the improved templating activity of the deletion derivative. This and related ribozymes were able to act as both ligase and template in the assembly of up to 18 oligonucleotides into a full length strand, complementary to the ribozyme sequence.

Despite these initial positive results, further attempts to evolve this ribozyme failed to yield similarly dramatic improvements. Efforts to evolve tighter binding of the primer–template complex, and to eliminate its preference for a wobble base pair at the ligation junction (a major source of infidelity) resulted in only small improvements (A.J.H., K. Chapman and J.W.S., unpublished results), while other even more intractable problems loomed ahead, for example, the energetic neutrality of the transesterification reaction and the inability of the ribozyme to perform efficient mononucleotide addition.

### *In vitro* selection of replicase candidates from random pools

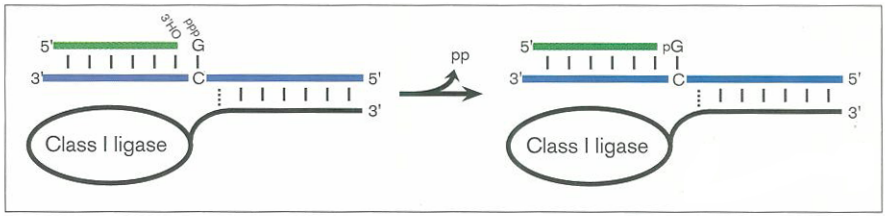
Due in part to the limitations of the group I reaction and in part to dramatic successes in the application of *in vitro* selection to the isolation of RNAs of defined function from complex pools of random sequences (for a review, see [4]), a new approach to the search for an RNA replicase has taken favor. The most promising path to viable replicase candidates now appears to be the direct selection of ribozymes with polymerase-related catalytic activities from large random or nearly random RNA sequence pools of high complexity (~$10^{15}$–$10^{16}$ independent sequences). RNA pools of this complexity represent only a tiny fraction of all possible sequences (sequence space) for strands 50–300 nucleotides in length, so the ribozymes isolated from these pools tend to be extremely sub-optimal sequences. However, we have found that the activity of these 'primary' ribozymes can be greatly improved by a combination of further *in vitro* selection and sequence design. This approach has been very fruitful in the isolation of new ribozyme ligases, and their evolution into polymerases that are potential replicase candidates.

Starting with a pool of approximately $10^{15}$ different transcripts containing a core of 220 random bases, Bartel and coworkers [18] carried out 10 rounds of iterative *in vitro* selection for RNAs capable of ligating an RNA oligonucleotide to their own 5' end. As each transcript began with a 5'-triphosphate, the ligation reaction was designed to be chemically analogous to the reaction catalyzed by modern polymerases, with pyrophosphate serving as the leaving group. After detection of catalytic activity in the pool molecules following three rounds of selection, mutagenic PCR in combination with increasingly stringent selection led to the evolutionary optimization of this activity to a ligation rate of $8\ h^{-1}$. From this evolved pool, three distinct structural classes of ribozyme ligases have been isolated and characterized [19]. The most interesting of these isolates is the Class I ligase, the only one of the isolated ribozymes that generates a 3'-5' phosphodiester linkage in the ligated reaction product. The Class I ligase is a complex ribozyme that is almost 100 nucleotides long. Further evolutionary optimization of the class I ligase [20] led to an extremely efficient variant with a $k_{cat}$ of greater than $1\ s^{-1}$ [19], a level of activity comparable to that of the protein enzyme DNA ligase.

In a dramatic recent development, Ekland and Bartel [21] have shown that the optimized Class I ligase can act as a nascent RNA polymerase, extending an RNA primer bound to a separate template strand by one nucleotide in a template-directed manner using nucleoside triphosphates as substrates (Fig. 3). When the template was covalently linked to the ribozyme, the primer could be extended by up to three nucleotides, and when the primer was designed to be able to slip on the template, primer extension by up to six nucleotides was observed [21]. The mononucleotide addition reaction performed by the Class I ligase is chemically identical to that performed by the protein enzymes RNA polymerase and DNA polymerase. The major limitation on the polymerase activity of the Class I ligase is that the ribozyme binds the template strand largely through base-pairing interactions. Once the primer has been extended up to the border of the template–ribozyme duplex, polymerization stops. Remarkably, the Class I ligase is not only capable of using all four trinucleotides as substrates, but it does so with reasonable fidelity [21]. The calculated fidelity of the extension reaction is 84 % in the

**Figure 3**

Single nucleotide addition reaction catalyzed by the optimized Class I ligase. Figure adapted from [21].




presence of equal concentrations of the four NTPs; lowering the GTP concentration 10-fold increased the calculated fidelity to 92 %. A further improvement of only 10-fold (i.e. an error rate of < 1 %) would lead to a polymerase sufficiently accurate to copy its own sequence.

It remains to be seen whether the Class I ligase can be evolved all the way to a replicase. Despite the remarkable progress made with this 'unnatural' ribozyme, many challenges remain. Chief among these is clearly the requirement for a non-sequence-specific mode of template binding. A domain that binds the template backbone in a non-sequence-specific manner might be sufficient to convert this ribozyme into a complete polymerase. Additional problems that must be overcome include low NTP affinity that is strongly biased towards GTP and weak, yet sequence-specific interactions with the primer–template duplex. These problems can, for the most part, be directly attributed to the conditions used in the original selection. Additional *in vitro* evolution may well overcome these weaknesses of the Class I ligase.

### Future replicase selection possibilities

Even if the Class I ligase cannot ultimately be evolved into a replicase, many other possibilities exist. Many different RNA sequences capable of catalyzing ligation were isolated in the original ligase selection, and some of these may prove to be interesting candidates. Independent selections would be likely to yield additional ligases, some of which might be as good as or better than the Class I ligase. Indeed, an analysis of the informational complexity of the Class I ligase, and the low probability of recovering such a complex ribozyme from the starting pool, suggests that there may be many ribozymes in sequence space that are as complex, and perhaps as active, as the Class I ligase.

Although work with the Class I ligase has focused on producing a ribozyme that mimics modern RNA synthesis by the polymerization of nucleoside triphosphates, efforts to evolve a replicase need not be limited to such an approach. Alternatively-activated nucleotides may offer some distinct advantages and possibly even historical relevance. For instance, AppN 5'-5' dinucleotides accumulate in reaction mixtures of highly activated nucleotides, and may have been important in prebiotic nucleic acid synthesis [22]. The adenosine 'handle' on the AMP leaving group may help facilitate ribozyme interactions, thereby aiding catalysis. Recently, ribozymes that catalyze the ligation of RNA to RNA 'capped' with an adenosine 5'-5' pyrophosphate have been isolated (A.J.H. and J.W.S., unpublished results). Other selections for ribozyme and deoxyribozyme ligases that use imidazole as an activating group have been successful [23,24]. With the use of highly activated mononucleotide substrates, it may be both possible and preferable to select directly for mononucleotide addition.

## Possible role of RNA in protein synthesis

Since the discovery of catalytic ribonucleic acids (RNA) [25,26], the thought that nucleic acids play much more than a structural role in the ribosome has been greatly bolstered. Studies on the mechanism of resistance to protein synthesis inhibitors provided the first evidence to support this hypothesis, as resistance often resulted from changes in the ribosomal RNA (rRNA) and not the ribosomal proteins [27]. Noller *et al.* [28] have experimentally addressed the notion that RNA may be responsible for the peptidyl transferase activity of the modern ribosome by showing that ribosomes are still active in peptidyl transfer following vigorous protein extraction procedures. Although this result strongly suggests that 23S rRNA is the ribosomal peptidyl transferase, residual protein in the extracted RNA has precluded the unambiguous assignment of the catalytic activity to the RNA component of the ribosome.

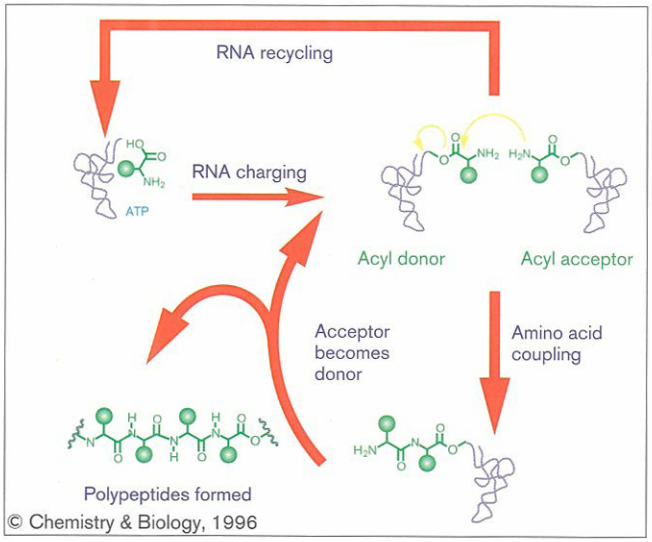
Recently, a tenuous but intriguing connection between group I introns and ribosomal RNA has arisen through the finding that aminoglycoside antibiotics that block protein synthesis also inhibit the splicing activity of group I introns [29]. This finding coupled to the discovery that the group I intron also possesses a (weak) aminoacyl esterase activity [30] suggested that the self-splicing intron and the ribosome may share common structural as well as functional motifs. The possibility of shared functional composition is also supported by the finding that a group I ribozyme mutant optimized through *in vitro* evolution for phosphodiester transfer reactions on a DNA substrate can also catalyze amide-bond cleavage [31], albeit more slowly than originally thought ($k_{cat}$ of ~$10^{-7}$ $min^{-1}$, a rate acceleration of some 100-fold over the estimated background rate [32]).

Noller and others [33,34] have suggested that primitive ribosomes may have evolved at least in part from an ancestral RNA related to the group I intron that was able to catalyze the acyltransfer reaction needed in peptide-bond formation . The attachment of amino acids or peptides to primitive ribozymes may have increased their catalytic efficiency, providing a driving force for the further evolution of peptide synthesis. Taken together these studies imply that modern rRNA may have a catalytic role in protein synthesis and that some contemporary RNAs may be evolutionary fossils of the RNA world.

### Primitive peptide synthesis

Peptide synthesis is not chemically difficult, given appropriately activated amino acids. Various activated amino acids will condense even in aqueous solution, yielding short peptides, and peptides over 50 amino acids long can be formed on mineral surfaces [7]. Furthermore, a simple template system has recently been described, in which a thioester-activated peptide spontaneously transfers from an RNA donor to an oligodeoxynucleotide acceptor containing a 3' terminal amine [35]. If montmorillonite or simple templates can catalyze peptide or amide synthesis, surely a ribozyme could do as much or more? The attraction of looking for ribozyme catalysts of peptide synthesis is not so much the difficulty of the chemical transformation, but the quest for specificity, and the simplicity with which RNA templating could have led to coded protein synthesis. Primitive RNAs that could charge themselves with activated amino acids (Fig. 4) could potentially function as both donors and acceptors in subsequent ribozyme-catalyzed acyl-transfer reactions. Five roles for RNA must be demonstrated experimentally to show that modern translation could have evolved in the RNA world: amino-acid recognition, RNA aminoacylation, acyl-transfer reactions, amino-acid activation and peptide synthesis. The first three of these steps have now been accomplished, and the last two are unlikely to remain hypothetical for long.

**Figure 4**




Hypothetical model for a primitive RNA-based translation system. Ribozymes could charge themselves with amino-acid adenylates; these charged RNAs may have used base pairing to a cofactor RNA to augment the specificity of the coupling reaction between donors and acceptors. Following amino-acid coupling to the acceptor, uncharged donors are then recycled to complete the peptide-bond-forming cycle.

### RNA can recognize amino acids

Two amino-acid-binding RNA motifs have been found in nature: the group I intron fortuitously binds arginine at a site within the P7 helix, which overlaps with the guanosine binding site [36], and the transactivating region (TAR) of HIV mRNA, which specifically binds the transactivator protein tat [37,38], also binds free arginine [39]. Furthermore, it has become evident that the editing reactions of the aminoacyl transfer RNA (tRNA) synthetases involve RNA-dependent amino-acid recognition [40], which may (or may not) involve direct RNA–amino acid contacts.

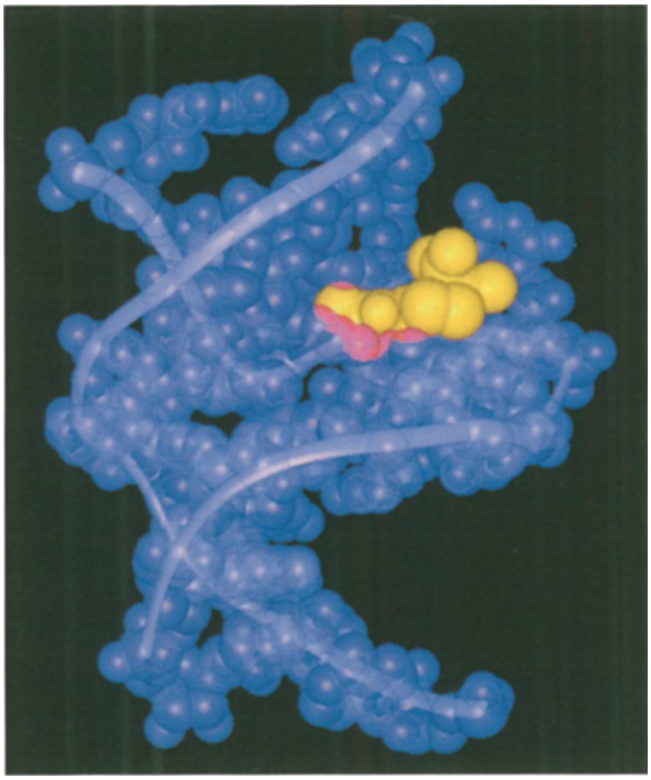
*In vitro* selection of RNA aptamers from random sequence pools has yielded additional amino acid recognizing RNA structures. Four different arginine-binding motifs [41–44] with varying degrees of specificity have been described. The structural basis of many of these aptamers is starting to come into focus, and Tang *et al.* [45] have recently solved by NMR the solution structure of an arginine aptamer and a variant of this aptamer that recognizes citrulline (Fig. 5). RNAs that recognize more hydrophobic amino acids such as valine [46] and tryptophan [47] have also been isolated by *in vitro* selection.

### RNA self-aminoacylation

Protein aminoacyl-tRNA synthetases facilitate the formation of 2'(3') aminoacylated RNA though a two step process that involves the activation of the amino acid by ATP to form a 5'-aminoacyl adenylate that is subsequently attacked by the 2'(3') end of a tRNA to yield a charged amino-acid carrier. This charged tRNA is then a substrate for ribosome-mediated peptide-bond formation. A ribozyme has recently been isolated that performs a reaction that mimics the second of these steps — charging of RNA with an activated amino acid.

Yarus and colleagues [48] searched for self-aminoacylating RNAs by *in vitro* selection, starting with a pool of $1.7 \times 10^{14}$ different RNA molecules, each containing 50 random positions. The pool RNA was incubated with chemically synthesized phenylalanyl-5'-adenylate at 0 °C in the presence of $Mg^{2+}$ and $Ca^{2+}$. RNAs that accelerated their own aminoacylation were selected by allowing the free α-amino group of phenylalanine to react with the N-hydroxysuccinimide ester of napthoxyacetic acid. The resulting hydrophobically labeled RNAs were separated from the more hydrophilic pool by reverse phase HPLC. Following 11 cycles of

**Figure 5**



The arginine aptamer (blue) folds around the amino acid making arginine (yellow) an integral part of the overall structure. Figure adapted from [45].

selection, a series of self-aminoacylating RNAs was isolated, including one with a rate acceleration of 105 over the extrapolated background rate (Fig. 6). The reactive RNAs required both $Mg^{2+}$ and $Ca^{2+}$; a role for $Ca^{2+}$ in the coordination of the 2'-3' cis-diol of the RNA has been suggested. Amino-acid recognition appears to be quite weak, as this ribozyme did not exhibit saturable substrate binding.

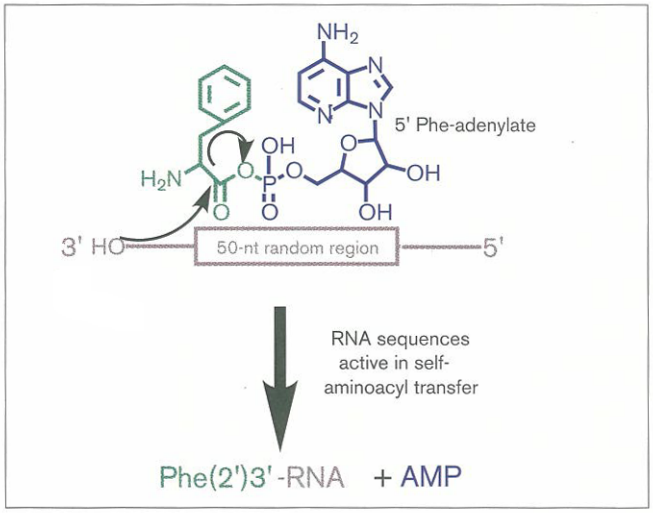
### RNA-catalyzed acyl transfer

The universal involvement of aminoacyl adenylates and aminoacylated RNAs in the translation apparatus almost certainly reflects the origins of translation in early RNA-mediated acyl-transfer reactions. Ribozymes that could mediate acyl transfer between donor and acceptor oligonucleotides would have increased the flexibility of the developing RNA world metabolism by allowing a self-aminoacylating RNA to hand off its amino acid to RNAs specialized for other tasks such as further metabolic transformations. Such acyl-transferases could have been the evolutionary precursors of the most important acyl-transferase of all, the peptidyl transferase of the ribosome. The recent progress in the isolation of aptamers and ribozymes with functions related to those needed for protein synthesis inspired us to search directly for ribozymes that could accelerate acyl-transfer reactions.

A simplified version of the ribosomal peptidyl transferase reaction known as the fragment reaction, in which a hexanucleotide fragment of charged $tRNA^{met}$ functions as an acyl donor, has been widely used to study catalysis by the ribosome [49]. Lohse and Szostak [50] have used a variant of the fragment reaction as the basis of an *in vitro* selection scheme for the isolation of acyl-transferase ribozymes (Fig. 7). The acyl transferases were isolated by incubating a pool of ~$10^{15}$ different RNAs, each with 90 random positions, with an acyl-donor substrate consisting of a hexanucleotide charged at its 2'(3') end with N-biotinylated methionine. Calf-intestinal alkaline phosphatase was used to remove the 5'-triphosphate from the pool RNA, exposing the desired 5'-hydroxyl acyl acceptor. Following incubation of the pool and acyl donor, sequences that accelerated transfer of the N-biotinylated methionine to their own 5'-hydroxyl group were selected on streptavidin agarose. Eleven cycles of *in vitro* selection and evolution yielded a final pool dominated by one class of ribozymes.

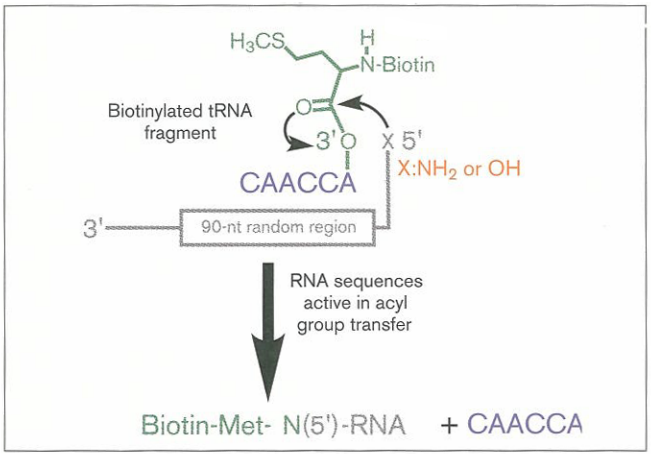
Sequence analysis revealed that a highly conserved internal template, located near the 3' end of the ribozyme, is available to bring the 2'(3') end of the donor and the 5' end of the acceptor together. The ribozyme does more, however, than simply act as a template: its $k_{cat}$ is $9.4 \times 10^{-2}$ $min^{-1}$, a rate acceleration of $10^3$ over the rate observed for acyl-transfer in a template-only oligonucleotide model system. Surprisingly, two G:U wobble base pairs at the junction of the donor/template duplex and the acceptor/template duplex are important in the reaction. The template-only background reaction is optimal with an all Watson-Crick duplex; the two G:U base-pairs in this system cause a 10-fold rate decrease, while in the ribozyme, an all Watson Crick duplex is 10-fold less effective

**Figure 6**




RNA can mediate aminoacyl-RNA synthesis.

**Figure 7**




RNA can mediate aminoacyl-transfer reactions. Figure adapted from [50].

than the selected double wobble base-pair duplex. This ribozyme, like almost all other ribozymes studied to date, is dependent on divalent cations for activity. We are currently exploring the role of the wobble base pairs and the divalent metal ion in catalysis.

### Ribozyme-mediated amide-bond formation

The ribosome is able to catalyze aminoacyl transfer onto either hydroxyl or amino acceptors, thus generating ester- or amide-linked products [49]. To see if the ribozyme described above, which had been evolved to accelerate a carboxy-ester transesterification reaction, could also catalyze amide-bond synthesis, the 5'-hydroxyl group of the RNA was replaced with an amino group [50]. This was done by transcribing the RNA in the presence of 5'-amino, 5'-deoxy guanosine, which can act as an initiator of transcription by T7 RNA polymerase. The modified ribozyme was able to transfer the biotinylated methionine from the 2'(3') end of the substrate hexanucleotide to its own 5' amino group almost as efficiently as it had transferred the amino acid to its 5' hydroxyl group (Fig. 7). The first order rate constant for amide-bond formation (0.58 $min^{-1}$) is only 15-fold slower than that of the peptidyl-transferase activity in the ribosomal fragment reaction (8 $min^{-1}$).

### RNA bridges to modern translation

The isolation of aptamers that can recognize specific amino acids, and of ribozymes that accelerate self-aminoacylation and acyl-transfer reactions, strongly supports an RNA-centric view of the origin of translation. Given that RNAs can mobilize the chemical energy of the phosphoanhydride bonds of ATP [51] it seems likely that RNAs will be found that can activate amino acids by catalyzing the synthesis of aminoacyl adenylates. Similarly, the example of ribozyme-catalyzed amide-bond synthesis suggests that it will not be difficult to demonstrate RNA-catalyzed peptide synthesis.

If all of the catalytic steps of protein synthesis can be mediated by ribozymes, what about the evolution of coded protein synthesis? The random assembly of peptides could have been catalyzed by a ribozyme with little or no specificity for its aminoacyl donor and acceptor substrates. Specific peptides could then have been generated by ribozymes that recognized a particular donor and a particular acceptor. This is a cumbersome mechanism for peptide synthesis that requires a different ribozyme for each linkage in a given peptide (however, many small peptides are synthesized in just this manner by protein enzymes in bacteria). The activity of such a specific ribozyme could have been augmented by the presence of a cofactor RNA which bound to both the donor and acceptor aminoacylated RNAs by base-pairing, helping to bring them together. Once such cofactor RNAs evolved, there would be less need for specificity to reside in the ribozyme itself; ultimately, a general purpose peptidyl transferase (rRNA) could evolve, with all specificity transferred to the template (mRNA) that directs the successive pairing of donor and acceptor RNAs (tRNAs).

### RNA sequence space and the origin and early evolution of life

Studies of modern ribozymes and ribonucleoprotein complexes continue to expand the known role of RNA in cellular processes. The finding of a catalytic role for RNA in translation and nuclear mRNA splicing would further support a critical role for RNA catalysis in the early evolution of life. However, the vast majority of biological enzymes are now protein enzymes, and the limited selection of ribozymes found in nature all act on RNA substrates. It therefore seems likely that the exploration of sequence space by *in vitro* selection and evolution may be a more fruitful route to establishing the scope of the catalytic ability of RNA (including modified RNA and related nucleic acids). *In vitro* selections typically sample only a small fraction of total RNA sequence space, yet numerous ribozymes, facilitating reactions critical to both RNA replication and translation, have been found. The challenge of completing the laboratory re-creation of these two aspects of life in the RNA world is clear.

## References

1. Gilbert, W. (1986). The RNA world. *Nature* **319**, 618.
2. Ellington, A.D. & Szostak, J.W. (1990). *In vitro* selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822.
3. Tuerk, C. & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–510.
4. Lorsch, J.R. & Szostak, J.W. (1996). Chance and necessity in the selection of nucleic acid catalysts. *Accounts Chem. Res.* **29**, 103–110.
5. James, K. & Ellington, A. (1995). The search for missing links between self-replicating nucleic acids and the RNA world. *Orig. Life Evol. Biosph.* **25**, 515–530.
6. Lazcano, A. & Miller, S.L. (1996). The origin and early evolution of life: prebiotic chemistry, the pre-RNA world, and time. *Cell* **85**, 793–798.
7. Ferris, J., Hill, A., Jr., Liu, R. & Orgel, L. (1996). Synthesis of long prebiotic oligomers on mineral surfaces. *Nature* **381**, 59–61.

8. Ertem, G. & Ferris, J. (1996). Synthesis of RNA oligomers on heterogeneous templates. *Nature* **379**, 238–240.
9. Eigen, M. (1971). Self organization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**, 465–523.
10. Zaug, A.J. & Cech, T.R. (1986). The intervening sequence RNA of *Tetrahymena* is an enzyme. *Science* **231**, 470–475.
11. Been, M. & Cech, T. (1988). RNA as an RNA polymerase: net elongation of an RNA primer catalyzed by the *Tetrahymena* ribozyme. *Science* **239**, 1412–1416.
12. Bartel, D.P., Doudna, J.A., Usman, N. & Szostak, J.W. (1991). Template-directed primer extension catalyzed by the *Tetrahymena* ribozyme. *Mol. Cell. Biol* **11**, 3390–3394.
13. Doudna, J.A. & Szostak, J.W. (1989). RNA-catalyzed synthesis of complementary-strand RNA. *Nature* **339**, 519–524.
14. Doudna, J.A., Couture, S. & Szostak, J.W. (1991). A multisubunit ribozyme that is a catalyst of and template for complementary strand RNA-synthesis. *Science* **251**, 1605–1608.
15. Doudna, J.A., Usman, N. & Szostak, J.W. (1993). Ribozyme-catalyzed extension by trinucleotides: a model for the RNA-catalyzed replication of RNA. *Biochemistry* **32**, 2111–2115.
16. Chowrira, B.M., Berzal-Herranz, A. & Burke, J.M. (1993). Novel RNA polymerization reaction catalyzed by a group I ribozyme. *EMBO J.* **12**, 3599–3605.
17. Green, R. & Szostak, J.W. (1992). Selection of a ribozyme that functions as a superior template in a self-copying reaction. *Science* **258**, 1910–1915.
18. Bartel, D.P. & Szostak, J.W. (1993). Isolation of new ribozymes from a large pool of random sequences. *Science* **261**, 1411–1418.
19. Ekland, E.H., Szostak, J.W. & Bartel, D.P. (1995). Structurally complex and highly active ligases derived from random RNA sequences. *Science* **269**, 364–370.
20. Ekland, E.H. & Bartel, D.P. (1995). The secondary structure and sequence optimization of an RNA ligase ribozyme. *Nucleic Acids Res.* **23**, 3231–3238.
21. Ekland, E.H. & Bartel, D.P. (1996). RNA-catalyzed RNA polymerization using nucleoside triphosphates. *Nature* **382**, 373–376.
22. Orgel, L.E. (1986). RNA catalysis and the origins of life. *J. Theor. Biol.* **123**, 127–149.
23. Cuenoud, B. & Szostak, J.W. (1995). A DNA metalloenzyme with DNA ligase activity. *Nature* **375**, 611–614.
24. Chapman, K.B. & Szostak, J.W. (1995). Isolation of a ribozyme with 5'-5' ligase activity. *Chemistry & Biology* **2**, 325–333.
25. Cech, T.R., Zaug, A.J. & Grabowski, P.J. (1981). *In vitro* splicing of the ribosomal RNA precursor of *Tetrahymena*: involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell* **27**, 487–496.
26. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**, 849–857.
27. Cundliffe, E. (1981) Antibiotic inhibitor of ribosome function. In *The Molecular Basis of Antibiotic Action* (2nd edn), (Gale, E.F., *et al.*, eds), pp. 402–547, Wiley, New York.
28. Noller, H.F., Hoffarth, V. & Zimniak, L. (1992). Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* **256**, 1420–1424.
29. von Ahsen, U., Davies, J. & Schroeder, R. (1991). Antibiotic inhibition of group I ribozyme function. *Nature* **353**, 368–370.
30. Piccirilli, J.A., McConnell, T.S., Zaug, A.J., Noller, H.F. & Cech, T.R. (1992). Aminoacyl esterase activity of the *Tetrahymena* ribozyme. *Science* **256**, 1420–1424.
31. Dai, X., de Mesmaeker, A. & Joyce, G.F. (1995). Cleavage of an amide bond by a ribozyme. *Science* **267**, 237–240.
32. Dai, X., de Mesmaeker, A. & Joyce, G.F. (1996). Amide cleavage by a ribozyme: correction. *Science* **272**, 18–19.
33. Noller, H.F. (1991). Ribosomes. Drugs and the RNA world. *Nature* **353**, 302–303.
34. Werner, C., Krebs, B., Keith, G. & Dirheimer, G. (1976). Specific cleavages of pure tRNAs by plumbous ions. *Biochim. Biophys. Acta* **432**, 161–175.
35. Bruick, R.K., Dawson, P.E., Kent, S.B.H., Usman, N. & Joyce, G.F. (1996). Template-directed ligation of peptides to oligonucleotides. *Chemistry & Biology* **3**, 49–56.
36. Michel, F., Hanna, M., Green, R., Bartel, D.P. & Szostak, J.W. (1989). The guanosine binding site of the *Tetrahymena* ribozyme. *Nature* **342**, 391–395.
37. Weeks, K.M., Ampe, C., Schultz, S.C., Steitz, T.A. & Crothers, D.M. (1990). Fragments of the HIV tat protein specifically bind TAR RNA. *Science* **249**, 1281–1285.
38. Calnan, B.J., Tidor, B., Biancalana, S., Hudson, D. & Frankel, A.D. (1991). Arginine-mediated RNA recognition: the arginine fork. *Science* **252**, 1167–1171.
39. Tao, J. & Frankel, A.D. (1992). Specific binding of arginine to TAR RNA. *Proc. Natl. Acad. Sci. USA* **89**, 2723–2726.
40. Schmidt, E. & Schimmel, P. (1995). Making connections: RNA-dependent amino acid recognition. *Trends Biochem. Sci.* **20**, 1–2.
41. Connell, G.J., Illangesekare, M. & Yarus, M. (1993). Three small ribooligonucelotides with specific arginine sites. *Biochemistry* **32**, 5497–5502.
42. Famulok, M. (1994). Molecular recognition of amino acids by RNA aptamers: an L-citrulline-binding RNA motif and its evolution into an L-arginine binder. *J. Am. Chem Soc.* **116**, 1698–1706.
43. Geiger, A., Burgstaller, P., von der Eltz, H., Roeder, A. & Famulok, M. (1996). RNA aptamers that bind L-arginine with sub-micromolar dissociation constants and high enantioselectivity. *Nucleic Acids Res.* **24**, 1029–1036.
44. Tao, J. & Frankel, A.D. (1996). Arginine-binding RNAs resembling TAR identified by *in vitro* selection. *Biochemistry* **35**, 2229–2238.
45. Tang, Y., Kochoyan, M., Burgstaller, P., Westhof, E. & Famulok, M. (1996). Structural basis of ligand discrimination by two related RNA aptamers resolved by NMR spectroscopy. *Science* **272**, 1343–1347.
46. Majerfeld, I. & Yarus, M. (1994). An RNA pocket for an aliphatic hydrophobe. *Nat. Struct. Biol.* **1**, 287–292.
47. Famulok, M. & Szostak, J.W. (1992). Stereospecific recognition of tryptophan agarose by *in vitro* selected RNA. *J. Am. Chem. Soc.* **114**, 3990–3991.
48. Illangasekare, M., Sanchez, G., Nickles, T. & Yarus, M. (1995). Aminoacyl-RNA synthesis catalyzed by an RNA. *Science* **267**, 643–647.
49. Fahnestock, S., Neumann, H., Shashoua, V. & Rich, A. (1970). Ribosome-catalyzed ester formation. *Biochemistry* **9**, 2477–2483.
50. Lohse, P.A. & Szostak, J.W. (1996). Ribozyme-catalyzed amino-acid transfer reactions. *Nature* **381**, 442–444.
51. Lorsch, J.R. & Szostak, J.W. (1994). *In vitro* evolution of new ribozymes with polynucleotide kinase activity. *Nature* **371**, 31–36.