G. F. Joyce

DOI: 10.1002/anie.200701369

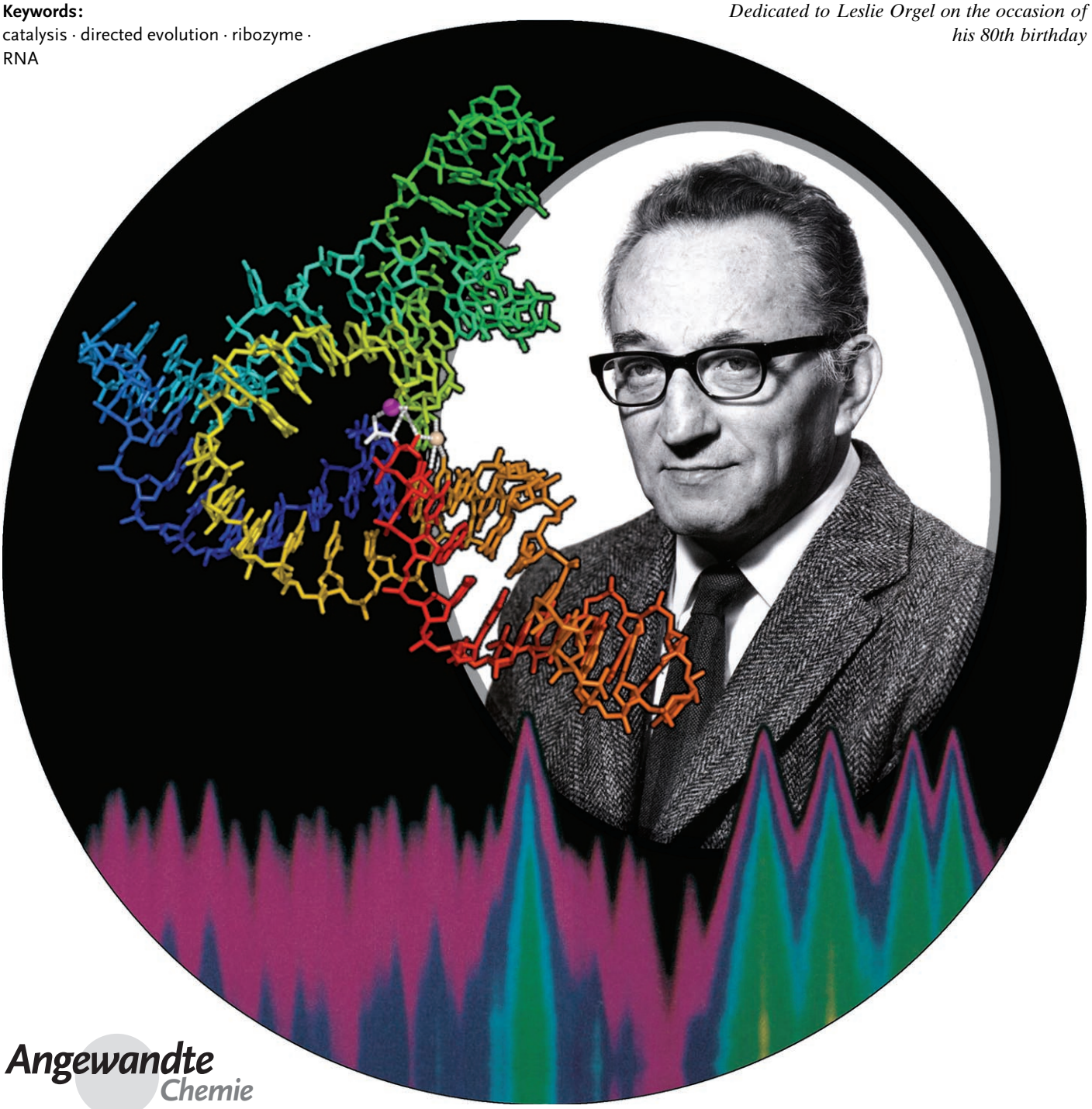
Molecular Evolution

Forty Years of In Vitro Evolution**

*Gerald F. Joyce**

Keywords:
catalysis · directed evolution · ribozyme · RNA

Dedicated to Leslie Orgel on the occasion of his 80th birthday



It has been 40 years since Spiegelman and co-workers demonstrated how RNA molecules can be evolved in the test tube. This result established Darwinian evolution as a chemical process and paved the way for the many directed evolution experiments that followed. Chemists can benefit from reflecting on Spiegelman's studies and the subsequent advances, which have taken the field to the brink of the generation of life itself in the laboratory. This Review summarizes the concepts and methods for the directed evolution of RNA molecules in vitro.

From the Contents

1. Introduction

In 2009 the scientific community will celebrate the 200th anniversary of the birth of Charles Darwin and the 150th anniversary of the publication of his book *On the Origin of Species by Means of Natural Selection*.[1] The impact of Darwin's studies on the biological sciences cannot be overstated. They have provided a unifying principle for biology that is central to the understanding of all biological form and function. As biological organization is fundamentally a chemical phenomenon, the principles of Darwinian evolution also apply at the level of molecules. Darwin died long before the macromolecular basis of biological evolution was understood, and could not have envisioned that the complete set of genetic instructions would be known for many organisms, including humans. Nor did he imagine that Darwinian evolution would be carried out as a purely chemical process, completely outside the context of biological systems.

This year the chemistry community also celebrates the 40th anniversary of the first in vitro evolution experiment. On July 15, 1967, Sol Spiegelman and co-workers published a study in the *Proceedings of the National Academy of Sciences USA* entitled "An Extracellular Darwinian Evolution Experiment with a Self-Duplicating Nucleic Acid Molecule".[2] Today Spiegelman's molecules would not be described as "self-duplicating", but his studies are recognized as being the beginning of what has become a highly fruitful area of investigation. His initial study and the many experiments that followed have had a large impact on both the chemical and biological sciences.

Darwinian evolution can be regarded as the concerted operation of three chemical processes that occur at the level of informational macromolecules: amplification, mutation, and selection (Figure 1). Amplification (genetic replication) is the process by which a number of "parent" molecules give rise to a larger number of "progeny" molecules. Normally one thinks of "parent" and "progeny" as applying to cells or whole organisms, but if the parent is a heteropolymer with a particular sequence of subunits, and if that molecule is duplicated to produce copies that have a very similar sequence of subunits, then there is a transfer of information

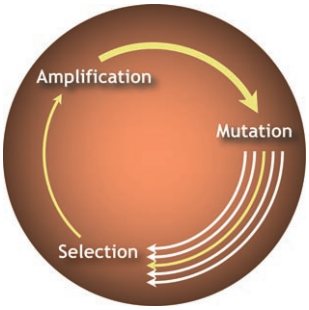


Figure 1. The Darwinian evolution of molecules entails the three processes of amplification, mutation, and selection. Amplification involves the copying of parent molecules to produce a larger number of progeny molecules. Mutation introduces variation among the population of progeny molecules. Selection chooses those molecules that meet constraints imposed by the environment. The selected molecules (yellow arrow) become the parents for the next round of evolution.

[*] Prof. G. F. Joyce
Departments of Chemistry and Molecular Biology
and
The Skaggs Institute for Chemical Biology
The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)
Fax: (+1) 858-784-2943
E-mail: gjoyce@scripps.edu

[**] The cover picture shows a photograph of Sol Spiegelman (1974; courtesy of the Albert and Mary Lasker Foundation) next to the crystal structure of an RNA enzyme that catalyzes the RNA-templated joining of RNA;[81] white: nucleotide at the ligation junction, purple sphere: Mg^{2+} ion thought to be involved in catalysis (courtesy of Michael Robertson and William Scott). The background shows a reconstructed space-time image of traveling waves of evolving Qβ RNA within a capillary tube;[21] horizontal axis: position along the capillary, vertical axis: fixed time intervals (from top to bottom), color gradation: concentration of RNA.

between successive molecules. As with cells or organisms, each molecule has a lifetime that depends on its intrinsic properties and the nature of its environment. Molecules that are amplified more rapidly than they undergo decomposition will persist in an environment, and the information they contain will be a record of a particular organization of subunits that allows such persistence.

Mutation (genetic variation), together with recombination, are processes by which differences arise between parent and progeny molecules. The copying of an informational macromolecule, even if it is brought about by sophisticated enzymatic machinery, results in occasional errors. Too many errors destroy the genetic information, but occasional errors result in subtle variation that may have advantageous consequences. If every copy were a perfect copy, then all of the molecules in a population would have identical chemical organization. So long as that particular organization allows persistence in the environment, the population will survive. But if the environment changes in such a way that the rate of amplification can no longer keep pace with the rate of decomposition, then the population of identical individuals will fall to extinction. Variation resulting from mutation provides an opportunity for long-term survival of the population of molecules because there is the possibility that some variants will have an organization of chemical subunits that allows their persistence in a changed environment. Furthermore, such variants can give rise to additional variants that may be even more suited to the changed environment.

The third process of Darwinian evolution is selection (phenotypic preference). The complete title of Darwin's 1859 book is *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*.[1] Viewed in molecular terms, the "struggle" is the requirement to amplify an informational macromolecule more rapidly than it undergoes decomposition, and to do so in competition for limited resources with other macromolecules in the local environment. The favored sequences are those that are amplified most rapidly and assuredly. Here amplification refers not just to the copying of genetic information, but to all processes that enable copying to occur. For biological organisms all of metabolism should be regarded as enabling processes. For macromolecules their intrinsic chemical properties and their interaction with other molecules are enabling attributes.

Darwin was well aware of techniques for the selective breeding of plants and domesticated animals, and how application of those techniques can lead to the artificial selection of desired traits. He sought to distinguish between artificial and natural selection to emphasize that the latter occurs without intentional direction toward a particular phenotypic trait, and leads not just to the emergence of new varieties but to entirely new species. It was more than a century after the acceptance of Darwin's ideas that artificial selection first was practiced on macromolecules in a non-biological context.[2] This could not have been achieved much earlier than Spiegelman's experiment in 1967, because, even though the principles of molecular evolution were already understood, the tools necessary to amplify, mutate, and select informational macromolecules were not yet available. In the 40 years since Spiegelman's experiment, those tools have become much more sophisticated and have allowed the practice of in vitro evolution to expand greatly and to enable one to select purposively for a variety of complex molecular traits.

Gerald Joyce studied at the University of California, San Diego, and received his PhD in 1984 for research carried out with Leslie Orgel at The Salk Institute. He joined the faculty of The Scripps Research Institute in 1989, where he is currently a Professor in the Departments of Chemistry and Molecular Biology. His research concerns the chemistry and biochemistry of RNA and the directed evolution of nucleic acid enzymes.

2. The First In Vitro Darwinian Evolution Experiment

In the mid-1960s Spiegelman and his colleagues, then at the University of Illinois, were studying RNA viruses to address the question of how viruses that have an RNA genome are replicated in cells that have a DNA genome. He reasoned that there would need to be a separate replication machinery for RNA, and in 1963 identified an RNA-dependent RNA polymerase (which he termed an "RNA replicase") that replicates the genomic RNA of MS-2 virus.[3] This virus is a bacteriophage that infects *Escherichia coli* cells, where it encounters many different RNA molecules within the host cell. The viral replicase only recognizes and copies MS-2 genomic RNA, and ignores the various bacterial RNA molecules. This "selective preference", as Spiegelman termed it, is critical for the efficient propagation of the virus.

In 1965 Spiegelman isolated the RNA replicase of a second bacteriophage named Qβ.[4] The Qβ replicase also was found to be specific for its corresponding genomic RNA, and the two replicases had no ability to replicate the other's genome. These studies were carried out entirely in vitro, using purified replicase protein, a small input of genomic RNA, and each of the four nucleoside 5′-triphosphates (NTPs). The input copies of genomic RNA were amplified by the replicase, and the new copies were amplified in a similar manner, which resulted in an exponential increase in the number of copies of RNA over a few hours.[5] After exponential amplification had been allowed to occur, Spiegelman purified the newly synthesized RNA molecules and showed that they retained the ability to form infectious viral particles.[6] This was the first synthesis of an infectious nucleic acid. For this study, as well as his contributions to the development of nucleic acid hybridization techniques, Spiegelman was awarded the Albert Lasker Basic Medical Research Award in 1974.

Spiegelman regarded the in vitro amplification of Qβ genomic RNA as a "self-duplication" process, but of course the RNA was not duplicating itself. Rather, the replicase protein was carrying out the duplication, and doing so in a

highly substrate-specific manner. We now know that this specificity derives from sequence elements and structural features of the genomic RNA that are recognized by the replicase protein.[7–9] The RNA itself has no catalytic function, but partners with the replicase to achieve exponential amplification. Recall that RNA replication (like DNA replication) requires the synthesis of both a "plus" and "minus" strand. The genomic RNA is the plus strand, which is copied to produce a complementary minus strand, and the minus strand in turn is copied to produce a new plus strand. Both the plus and minus strands must contain the requisite sequence and structural features that are recognized by the replicase protein.

In the mid-1960s little was known about the error rate of polymerases, although Spiegelman knew that Qβ replicase would produce occasional mutations and even worried that the mutation frequency might be too high "in the unfamiliar environment provided by the enzymologist".[6] Thus he recognized that all the pieces were in place to use the Qβ replication system to carry out the in vitro amplification, mutation, and selection of RNA molecules. His classic experiment began with a reaction mixture (250 μL) that contained 0.2 μg (0.16 pmol) of Qβ genomic RNA, 40 μg (0.19 nmol) of Qβ replicase, 0.8 mM of each NTP, 12.8 mM of $MgCl_2$, and 84 mM of Tris-HCl (pH 7.4; Tris = tris(hydroxymethyl)aminomethane), which was incubated at 35 °C for 20 min.[2] After 20 min had elapsed, an aliquot of 20 μL was taken from the mixture and transferred to a fresh reaction vessel that contained all of the reaction components except Qβ RNA. Only those RNA molecules that were carried over in the aliquot would have an opportunity to become amplified in the second reaction mixture. Because the aliquot represented only 8% of the original mixture, any RNA molecules that generated less than around a dozen copies during the first incubation would be reduced in copy number at the start of the second incubation. Conversely, the more copies generated from a particular RNA prior to transfer, the more opportunity that RNA would have to dominate the subsequent reaction mixture.

Spiegelman repeated the steps of incubation and transfer many times. This type of "serial-transfer" procedure is commonly used to culture living cells, but Spiegelman was using it to "culture" RNA molecules.[2] By the 9th transfer he noticed that the RNA molecules were being amplified substantially faster than at the outset, and beginning with the 14th transfer he decreased the incubation time from 20 min to 15 min. He further decreased the incubation time after the 29th, 38th, and 52nd transfers, to eventually reach just 5 min, which was continued through to the 74th transfer (Figure 2). By the end, the RNA molecules were being amplified 15-fold faster than at the beginning. This was largely due to the fact that the final molecules were only about 15% the size of the starting molecules, having jettisoned most of the nucleotides that are important for viral infectivity but are not necessary for RNA amplification in vitro. Qβ genomic RNA contains about 3600 nucleotides, while the "minivariants" that were obtained after 74 transfers contained only about 550 nucleotides.

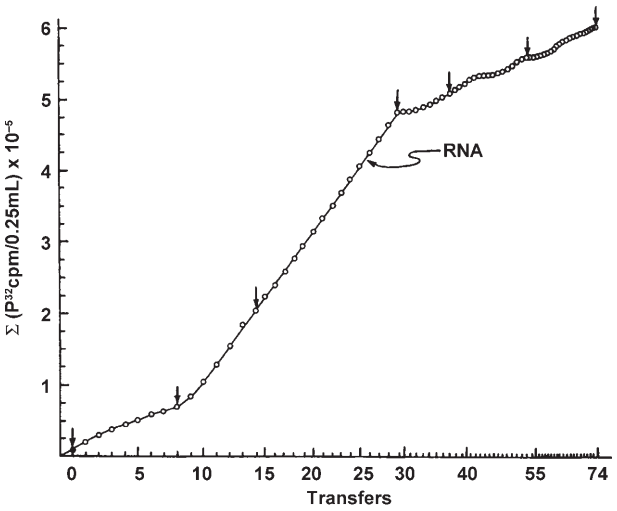


Figure 2. Spiegelman and co-workers carried out the in vitro evolution of Qβ RNA using a serial transfer procedure. This figure, reproduced from their original report in 1967,[2] depicts the accumulation of newly synthesized RNA during 74 successive transfers, as measured by the incorporation of [α-^{32}P]-labeled uridine 5′-triphosphate (UTP) into polynucleotides. Note that the growth rate increases at the 9th transfer (see arrow), and the time between transfers was reduced at the 14th, 29th, 38th, and 52nd transfers (see arrows).

Spiegelman described this result as unsurprising because he recognized that shorter RNA molecules would require less time for replication, and therefore would enjoy a selective advantage. However, he also remarked: "*It should not escape the attention of the reader that… other selective stresses can be imposed on the system to generate RNA entities which exaggerate other molecular features*".[2] In looking back at this first in vitro Darwinian evolution experiment some have diminished its significance by pointing out that the result was unsurprising and, given the stringent substrate selectivity of Qβ replicase, that no other outcome was even possible. In the modern era techniques are available for amplifying RNA molecules of almost any sequence, which allows selection pressure to be focused on the intrinsic properties of the RNA, rather than its ability to be a good substrate for a polymerase enzyme. Even in the Qβ evolution system, however, it is not simply a matter of the smaller molecules outcompeting the larger ones. It also is critical that the requisite sequence and structural recognition elements be maintained within both the plus and minus strands, even as substantial deletions occur elsewhere in the molecule.

Ironically, the first publication in the modern era (1990) that described the in vitro selection of an RNA "aptamer" also concerned an RNA that is recognized by a polymerase protein on the basis of particular sequence elements and structural features within the RNA.[10] Through the benefit of hindsight one might claim that it was Spiegelman who produced the first RNA aptamer, but he did not regard directed evolution as a general method with applications outside the context of viral replication. He was aware that in his experiments he was the first to realize the processes of Darwinian evolution in a nonbiological setting, and sometimes spoke immodestly of this accomplishment.[11] Immod-

esty aside, it is appropriate to compare Spiegelman's experiment with Wöhler's synthesis of urea.[12] Both organic chemistry and molecular evolution have progressed in ways that Wöhler and Spiegelman could not have imagined, and just as Wöhler broke the false distinction between biological chemistry and organic synthesis, Spiegelman broke the false distinction between Darwinian evolution in biology and Darwinian evolution as a chemical process.

3. Evolutionary Challenges for Qβ RNA

Spiegelman and others performed many fascinating experiments using the Qβ evolution system. Typically these experiments involved posing a new, often devious, challenge for the population of replicating RNA molecules, then allowing evolution to come up with a suitable answer. The first example with an added imposed selection pressure, other than the basic requirement for rapid replication, involved the progressive deprivation of cytidine 5′-triphosphate (CTP) from the reaction mixture.[13] The minivariants of Qβ RNA that Spiegelman had obtained previously require at least 100 μM of CTP for their efficient amplification. Accordingly, he reduced the concentration of CTP to 16 μM, carried out 10 rounds of serial transfer, then reduced the concentration of CTP to 8 μM and carried out an additional 40 transfers. At the end of this procedure the evolved molecules were amplified about 1.6-fold faster in the presence of 8 μM CTP compared with the starting molecules. The starting and evolved RNA molecules had the same size and overall nucleotide composition, but contained subtle differences in nucleotide sequence that accounted for the difference in their behavior.

Another evolution experiment involved the introduction of 7-deaza-ATP to the reaction mixture, which can be incorporated by Qβ replicase in place of ATP, but results in less efficient amplification.[13] The RNA molecules first were evolved for 16 transfers in the presence of a reduced concentration of ATP. Then evolution was continued for 19 transfers in the presence of 40 μM of ATP and 240 μM of 7-deaza-ATP. This resulted in an evolved variant that amplified twice as fast in the presence of 7-deaza-ATP compared with the starting RNA. Again, no substantial differences were seen in the size and overall nucleotide composition of the starting and evolved molecules.

3.1. The Molecular Evolution of Resistance

The same minivariant of Qβ RNA that was used to initiate the CTP-deprivation experiment also was used to initiate an in vitro evolution experiment in which progressively increasing concentrations of ethidium bromide (EtBr) were added to the reaction mixture.[14] EtBr intercalates into regions of double-stranded RNA and inhibits the activity of RNA polymerases, including Qβ replicase. Amplification of the starting Qβ RNA was reduced by 50% in the presence of 6 μM of EtBr and was inhibited completely in the presence of 60 μM of EtBr. In experiments carried out by Orgel and co-workers (at The Salk Institute in collaboration with Spiegelman),[14] the Qβ RNA first was evolved for 10 transfers in the presence of 5 μM of EtBr. Then the concentration of EtBr was increased to 10, 20, and so on up to 100 μM, which was continued through the 108th transfer. The final evolved molecules were amplified as efficiently in the presence of 80 μM of EtBr as the starting molecules in the presence of 6 μM of EtBr.

Further experiments that concerned the evolution of EtBr-resistant strains of Qβ RNA demonstrated that the evolved molecules are not simply faster replicators, but rather molecules that had specifically adapted to the presence of EtBr.[15] They bind EtBr less avidly, and in the absence of EtBr are amplified less efficiently compared with the starting RNA. The sequence of the EtBr-resistant RNA was determined and found to contain just three common mutations. Analysis of the population at intermediate times revealed that these mutations arose sequentially, with each successive mutation providing increased selective advantage. The mechanism by which the three mutations reduce the EtBr binding and relieve the inhibition of RNA amplification was not determined, although this likely could be ascertained using modern methods.

In the 1990s, thirty years after the first in vitro evolution experiment, an especially notable Qβ evolution experiment was carried out by Strunk and Ederhof,[16] in which they added ribonuclease A (RNase A) to the reaction mixture. This enzyme cleaves the phosphodiester linkage on the 3′-side of pyrimidine residues (cytosine and uracil) within unpaired regions of RNA. The minivariants of Qβ RNA contain roughly 50% pyrimidine residues in both the plus and minus strands, many of which occur in the loop and internal bulge regions and therefore are susceptible to cleavage by RNase A. Even if these regions contained only purine residues within one strand of Qβ RNA, the complementary strand would contain pyrimidine residues at the corresponding positions. Thus the evolution of resistance to RNase A presents a formidable challenge.

Amplification of the Qβ minivariant occurs at a reduced level in the presence of $1\,ng\,\mu l^{-1}$ of RNase A and is completely abolished in the presence of $2\,ng\,\mu l^{-1}$ of RNase A. A serial-transfer experiment was initiated in the presence of $1\,ng\,\mu l^{-1}$ of RNase A, and the concentration increased as tolerated until a final concentration of $3.5\,ng\,\mu l^{-1}$ of RNase A was reached.[16] Unlike the experiments of the 1960s and 1970s, this experiment was carried out using a serial-transfer machine that continuously monitored the concentration of RNA and triggered a transfer event whenever the concentration reached a predetermined threshold. In this way the evolving population was maintained under exponential-growth conditions and the population size was never allowed to fall below 10^{11} molecules. Both of these factors are important to ensure that the most advantageous individuals will grow to dominate the population.[17]

The evolution experiment was carried out for a total of 80 transfers, with an 11-fold dilution per transfer, corresponding to 280 generations of RNA replication.[16] The starting molecules contained 87 nucleotides in each strand, while the final evolved molecules contained only 65 (Figure 3). Remarkably, the solution to the imposed selection pressure

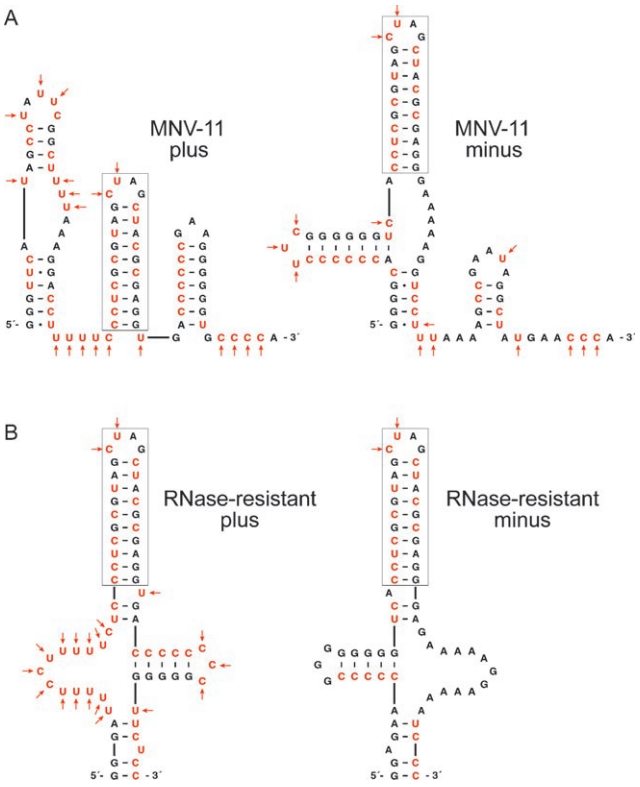


Figure 3. Evolution of a variant of Qβ RNA that is resistant to RNase A. A) A serial transfer experiment[16] was initiated with the "MNV-11" minivariant of Qβ RNA,[112] which contains roughly equal numbers of purine (black) and pyrimidine (red) residues within both the plus and minus strands. Most of the unpaired pyrimidine residues are susceptible to cleavage by RNase A (arrows). B) The evolved variant contains 20 susceptible pyrimidines in the plus strand, but only 2 in the minus strand; the latter occurs within a loop that is part of a structural feature (boxed region) that is recognized by Qβ replicase.

was to sacrifice the plus strand to preserve the minus strand. The evolved plus strand contained 42 pyrimidine residues, 20 of which were susceptible to RNase A cleavage, while the minus strand contained only 23 pyrimidine residues, only 2 of which were susceptible to cleavage. Consequently, the minus strand was maintained at much higher copy number, sustaining the population, while the plus strand was present as a short-lived replication intermediate.

3.2. Spatially-Isolated Evolving RNA Molecules

The evolution of Qβ RNA has been spatially isolated, either within molecular "colonies" that are grown on the surface of agarose[18] or within a long capillary tube that confines growth to a linear wavefront.[19] The molecular colony technique involves spreading the starting population of RNA over a thin layer of agarose that contains Qβ replicase, then covering the surface with a nylon membrane that is impregnated with the four NTPs. Each colony, which begins with a single RNA molecule, can give rise to as many as 10^{12} progeny RNA molecules. This technique has been used to demonstrate the occurrence of spontaneous recombination events between 5′ and 3′ fragments of Qβ RNA that must be combined to generate amplifiable material.[20]

The traveling-wave experiments employed a thin fluid-filled capillary that contained the Qβ amplification mixture plus 15 μM of EtBr. The EtBr was used to visualize the RNA molecules based on intercalation of EtBr into RNA, which results in enhanced fluorescence under UV light. In an early version of the system conventional photographs were taken,[19] but in a later version the entire capillary was monitored continuously using a CCD camera.[21] The capillary was threaded back and forth over a total length of 12 meters and RNA amplification was initiated at various points along its length. As copies of Qβ RNA were produced, a wave of fluorescence spread in both directions from the point of initiation, and the wavefront velocity provided a direct measurement of the rate of exponential RNA amplification. The wavefront typically moved at a velocity of 80 μm min^{-1}, which corresponds to an exponential growth rate of 0.7 min^{-1}.

One of the most satisfying aspects of the traveling-wave experiments is that they allow the phenotype for many hundreds of different wavefronts to be monitored simultaneously.[21] One can observe the distribution of amplification rates for a heterogeneous population of evolving RNAs. In some cases a slow-moving wave will suddenly increase in velocity as a new more advantageous variant arises. Secondary wavefronts can arise within the wake of a primary wave. If a secondary wave is driven by a more advantageous variant, then it can overtake the primary wave. Wavefronts can collide head-on, which may lead to the production of novel variants (presumably recombinants) that generate a secondary wavefront that travels faster than either of the two colliding waves.

Sol Spiegelman died in 1983, before Qβ evolution had been carried out in molecular colonies or traveling waves, and before the processes of in vitro evolution were generalized to include RNA molecules that are unrelated to Qβ RNA. In 1967 Spiegelman posed the question: "*What will happen to the RNA molecules if the only demand made on them is the Biblical injunction,* multiply*, with the biological proviso that they do so as rapidly as possible?*"[2] His experiment provided an initial answer to this question. Others have gone forth and multiplied upon his result, harvesting the fruits of in vitro evolution in ever more sophisticated ways.

4. And Then Came PCR

The high degree of substrate selectivity of Qβ replicase for RNA molecules that resemble Qβ genomic RNA makes sense from the perspective of viral evolution. It is highly advantageous for the replicase to ignore other RNA molecules that are present in the host cell, thereby focusing its activity on amplification of its own genome. For in vitro evolution studies, however, this extreme substrate selectivity is a severely limiting constraint. There are many interesting RNA molecules that do not conform to the substrate features required by Qβ replicase. Some attempts have been made to insert exogenous sequences into variants of Qβ RNA in the hope that those sequences would be carried along during the replication process.[22–24] Invariably, however, the added

sequences are discarded during in vitro evolution because they slow the rate of amplification.

The breakthrough for in vitro evolution came with the development of the polymerase chain reaction (PCR) and related methods for amplifying nucleic acids of almost any sequence.[25,26] This allowed RNA (and DNA) molecules to be selected based on properties other than their ability to be amplified by a replicase protein. PCR amplification of RNA requires reverse transcription of the RNA to a complementary DNA (cDNA), followed by PCR amplification at the level of DNA, and finally forward transcription of the DNA back to RNA. Another powerful but less frequently used method for nucleic acid amplification is isothermal RNA amplification.[27,28] This method relies on a combination of reverse transcriptase and DNA-dependent RNA polymerase to bring about repeated rounds of reverse and forward transcription to achieve RNA amplification at a constant temperature.

Both PCR and isothermal RNA amplification require the use of oligonucleotide primers to initiate strand synthesis, thus fixing the corresponding primer binding sites at the ends of the molecules that are being amplified. Reverse transcription, PCR amplification, and forward transcription all place some restrictions on the sequences that can be amplified with high efficiency. Some sequences cause the pausing or premature termination of the polymerase enzymes, and those sequences would be at a selective disadvantage in an in vitro evolution experiment. However, such disadvantaged sequences are rare compared to the vast number of well-tolerated sequences, in contrast to the highly restrictive situation with Qβ replicase and Qβ RNA.

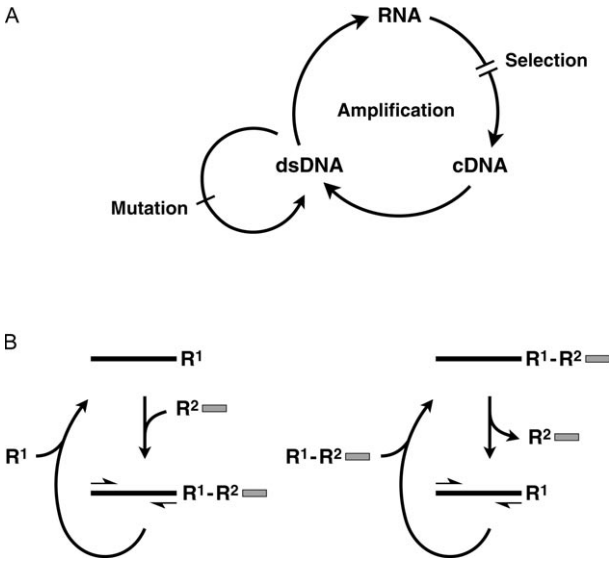


Figure 4. General representation of the in vitro evolution of functional RNA molecules. A) RNA is amplified by reverse transcription to complementary DNA (cDNA), conversion of the cDNA into double-stranded DNA (dsDNA), and transcription back to RNA. Additional amplification can be carried out at the level of dsDNA using the PCR. Mutations can be introduced during PCR amplification. Selection is imposed by the experimenter such that only those molecules that perform a chosen function are eligible for amplification. B) RNA molecules that catalyze either bond-forming (left) or bond-breaking (right) processes between the reactive groups R^1 and R^2 either acquire or lose a chemical tag (grey box), respectively, which is the basis for selection. Following amplification of the selected RNA molecules (small reciprocal arrows), the original terminus is restored to the progeny population.

4.1. Selection of RNA Aptamers

With a sequence-general method for nucleic acid amplification, the selection of an RNA phenotype can be separated from the amplification and mutation of a corresponding RNA genotype (Figure 4a). The experimenter is free to design a selection constraint that pertains to the intrinsic physical and chemical properties of the RNA. For example, the RNA can be selected for its ability to bind to a target ligand.[10,29] A heterogeneous population of RNA molecules can be exposed to an immobilized ligand, the unbound RNA molecules washed away, and the bound RNA molecules recovered and selectively amplified. Through repeated rounds of this procedure the population becomes progressively enriched with the RNA molecules that bind most strongly to the target. Such molecules are termed "aptamers" (from the Latin *aptus*, meaning "fit" or "suitable").[29] Often the term SELEX (systematic evolution of ligands by exponential enrichment) is used to refer to the process by which aptamers are obtained.[10]

The first reported examples of aptamers were small RNA molecules that bound to either T4 DNA polymerase[10] or organic dyes such as Cibacron Blue.[29] In the years that have followed, hundreds of RNA (and DNA) aptamers have been obtained that are specific for a broad range of protein and small-molecule targets.[30–32] Several RNA aptamers are being investigated as potential therapeutic agents.[33–37] One aptamer (pegaptanib, trade name Macugen), which binds tightly and specifically to vascular endothelial growth factor, has been approved for the treatment of age-related macular degeneration.[38,39]

4.2. Selection of Catalytic RNA Molecules

RNA molecules also can be selected on the basis of their catalytic function.[40] There are a handful of RNA enzymes (ribozymes) that have been discovered in nature, and a much larger number that have been obtained experimentally using in vitro evolution. Two general approaches have been established for the selection of catalytic RNA molecules that depends on whether they bring about a bond-forming or bond-breaking reaction (Figure 4b). Bond forming between two reactive groups R^1 and R^2 can be the basis of selection if one of the reactive groups (R^1) is attached to every RNA molecule in the population, and the other reactive group (R^2) is provided within a separate substrate. Any RNA molecule that brings about formation of a bond between R^1 and R^2 becomes modified in a way that is distinguishable from unreacted RNA molecules. R^2 typically is linked to a chemical tag, such as biotin or a short oligonucleotide, so that formation of the R^1–R^2 bond results in the attachment of

the tag to the RNA molecule that performed the reaction. The tagged RNA molecules are then isolated and selectively amplified. Conversely for a bond-breaking reaction, one begins with a population of RNA molecules that are attached to a tag through an R^1–R^2 linkage. Any molecules that catalyze the cleavage of the R^1–R^2 bond become detached from the tag. The untagged RNA molecules are then isolated and selectively amplified.

4.3. The Importance of Maintaining Population Diversity

In addition to methods for amplification and selection, in vitro evolution requires a method for introducing mutations. The polymerase enzymes that are used to amplify RNA have an intrinsic error rate of 10^{-4}–10^{-6} per nucleotide.[41–43] This is not sufficient to maintain sequence heterogeneity in a population of evolving RNA molecules that contain less than a few hundred nucleotides because almost every copy will be identical to its parent. The ideal error rate should result in an average of about one mutation per copy. This allows exploration of occasional higher-error mutants, but does not cause a runaway accumulation of mutations, which would prevent the most advantageous sequences from being enriched. There is an extensive theoretical and experimental literature concerning the optimal error rate for evolution and the consequences of exceeding the "error threshold", above which the ability to retain genetic information is lost.[44–47]

The most commonly used technique for introducing random mutations is error-prone PCR.[48,49] One well-defined protocol relies on altered reaction conditions and unbalanced concentrations of the four deoxynucleoside 5′-triphosphates (dNTPs) to achieve an error rate of 0.007 per nucleotide with no significant sequence bias.[50] Another protocol, termed "hypermutagenic PCR", uses more extreme reaction conditions and an increased number of temperature cycles to achieve an error rate of 0.1 per nucleotide, albeit with substantial sequence bias.[51] A somewhat different approach is to utilize an engineered thermostable DNA polymerase that has reduced fidelity, even under standard reaction conditions.[52,53] The Mutazyme enzyme (Stratagene Inc.), for example, gives an error rate of 0.001–0.007 per nucleotide, depending on the amount of input DNA.

An in vitro RNA evolution experiment usually begins with the synthesis of a heterogeneous population of RNA molecules. This can be done using error-prone PCR, but more commonly involves automated DNA synthesis employing solutions of each of the four nucleoside phosphoramidites that have been doped with a chosen amount of the other three nucleotide building blocks.[54] With this doping technique one can precisely control the sequence heterogeneity of the starting population at each nucleotide position. However, it is also important to introduce mutations during the course of evolution and across the entire population of molecules, which cannot be done by chemical synthesis. When additional mutations are introduced across the entire population they are distributed in proportion to the copy number of the various sequences that make up the population. The more advantageous sequences, which are present in higher copy number, automatically receive a larger proportion of the newly-introduced mutations, thus biasing the evolutionary search toward those sequences that have proven most advantageous in the past. There is no guarantee that the more advantageous sequences at present will give rise to even more advantageous ones in the future, but assuming that closely related sequences are likely to have similar phenotypes, it is the most efficient way to conduct the search.

5. A Phenotype Above All Others

Ribozymes have been obtained by in vitro evolution that catalyze a variety of bond-forming reactions, such as the formation of an RNA phosphodiester,[55] peptide bond,[56] glycosidic bond,[57] Diels–Alder-cycloaddition product,[58,59] Michael-addition product,[60] and aldol-condensation product.[61] Other ribozymes have been obtained that catalyze bond-breaking reactions, such as the cleavage of a DNA or RNA phosphoester,[62,63] carboxyester hydrolysis,[64] aminoacyl transfer,[65] and thiophosphate hydrolysis.[66] Each of these is worthy of an extended discussion, but in the context of this Review one reaction stands out above all others: the RNA-templated joining of RNA. Special attention is directed to the template-directed reaction between a 3′-hydroxy group of an oligonucleotide and a 5′-triphosphate of a nucleoside that results in the formation of a 3′,5′-phosphodiester and the release of an inorganic pyrophosphate. This is fundamentally the same reaction that is carried out by RNA polymerase proteins, including Qβ replicase. If a ribozyme could be obtained that catalyzes this reaction accurately and efficiently, then it might take the place of Qβ replicase (or a combination of polymerase proteins) in an in vitro evolution experiment. It would be especially intriguing if the RNA undergoing evolution was the RNA replicase ribozyme itself. Then in vitro evolution would operate in a self-sustained manner, with all of the informational macromolecules needed to bring about evolution being part of the system undergoing evolution.

The standard approach for obtaining a ribozyme with polymerase-like activity is to prepare a population of RNA molecules that bear a 5′-triphosphate, and attempt to react them with an oligonucleotide substrate that bears a 2′,3′-hydroxy group (Figure 5a). The ligated products can be selected in various ways; for example, based on their reduced mobility in a denaturing polyacrylamide gel or their acquisition of a tag sequence or biotin moiety contained within the attached substrate. To favor formation of a 3′,5′-phosphodiester, as opposed to a 2′,5′-phosphodiester, it has proven important to enforce Watson–Crick pairing surrounding the ligation junction[67] or to employ countermeasures that eliminate catalysts with 2′,5′-ligase activity.[68] The 3′,5′-linkage has special interest because it occurs in biological RNA molecules and is the linkage within the RNA molecules that are evolved to catalyze the ligation reaction. Once an RNA ligase ribozyme is obtained, it can be evolved to catalyze the templated addition of NTPs, ultimately leading to the development of an RNA-dependent RNA polymerase.

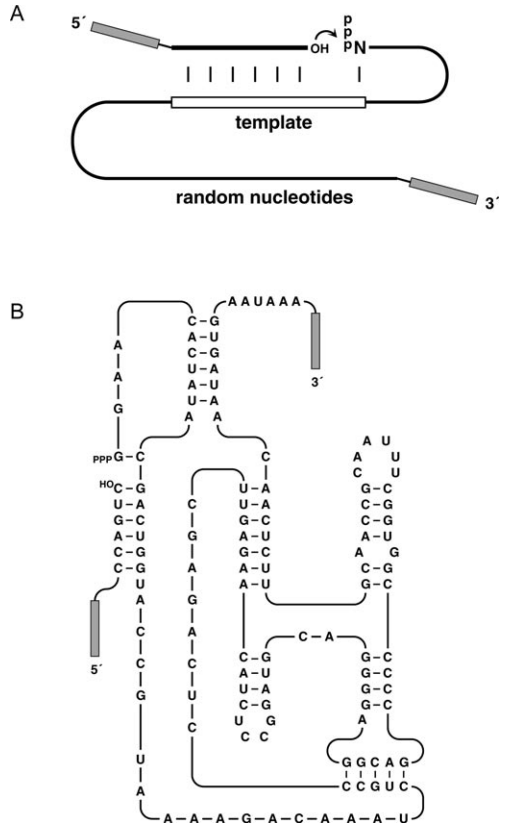


Figure 5. In vitro evolution of an RNA ligase ribozyme, starting with a population of random-sequence RNA molecules.[55] A) Selection is based on attack of a 3′-hydroxy group (contained within a template-bound oligonucleotide substrate) on the 5′-triphosphate of the population of RNA molecules. Only those RNA molecules that perform the reaction will contain both primer binding sites (grey boxes), which are necessary for reverse transcription and PCR amplification. B) Sequence and secondary structure of the b1-207 form of the class I ligase, which was obtained following 14 rounds of in vitro evolution.[69,70]

5.1. A Ribozyme that Joins RNA

The first in vitro evolution experiment that produced an RNA ligase ribozyme was carried out by Bartel and Szostak in 1993.[55] Actually, they isolated three different ligases, but only one of these, the so-called "class I" ligase, catalyzes formation of a 3′,5′-phosphodiester.[69] This ligase was obtained from a starting population of 10^{15} different RNA molecules, each of which bore a 5′-triphosphate and contained a central region of 220 random-sequence nucleotides. The challenge was to ligate an oligonucleotide substrate to the 5′-end of the molecules, and selection was based on hybridization of the products to a substrate-specific oligonucleotide affinity column and by PCR amplification using a primer that was specific for the attached substrate sequence. This selective amplification procedure was repeated for 10 rounds, with the reaction time progressively decreased from 16 h to 10 min. Random mutations were introduced using error-prone PCR following the fourth, fifth, and sixth rounds.

After the tenth round, individuals were cloned from the population, sequenced, and tested for catalytic activity. The most active catalyst that was identified was the class I ligase. It contains 186 nucleotides and has a catalytic rate of 0.029 min^{-1} (measured in the presence of 60 mM of $MgCl_2$ and 200 mM of KCl at pH 7.4 and 22 °C).[69] This molecule was rerandomized at a mutation frequency of 20 % per nucleotide position and subjected to 4 additional rounds of selective amplification. Based on the sequences of the final isolated molecules, various shortened forms of the class I ligase were designed and tested. One of these, designated the b1-207 ribozyme (Figure 5b), contains 119 nucleotides and has a catalytic rate of 14.4 min^{-1} (measured under the same conditions as above).[70] Under more extreme reaction conditions (for example, pH 9.0) and with some minor sequence modifications, a multiple-turnover rate of 360 min^{-1} can be achieved, the fastest ever reported for an RNA-catalyzed reaction.[71] The uncatalyzed rate of RNA-templated RNA ligation at pH 7.4 is 2×10^{-7} min^{-1} and at pH 9.0 is 1×10^{-6} min^{-1}.[72] Thus the catalytic rate enhancement of the b1-207 ribozyme is about 10^8-fold, which is comparable to that of many protein enzymes.

5.2. A Menagerie of Ligase Ribozymes

Five other ribozymes have been obtained by in vitro evolution that catalyze the RNA-templated joining of an oligonucleotide 3′-hydroxy group and an oligonucleotide 5′-triphosphate (Table 1). One of these, the hc ligase, was derived from 85 random-sequence nucleotides that were attached to a structural scaffold that occurs as an independent folding domain within a naturally-occurring ribozyme.[73] Following 10 rounds of selective amplification, with occasional error-prone PCR, a ribozyme was obtained that contains 337 nucleotides and has a catalytic rate of 0.26 min^{-1} (measured in the presence of 50 mM of $MgCl_2$ and 200 mM of KCl at pH 7.5 and 50 °C). This ribozyme was termed the hc ligase because it operates in a sequence-general manner within the helical context of a Watson–Crick duplex. Like the class I ligase, however, the template for the hc ligase is a region within the ribozyme rather than a separate molecule.

Subsequent in vitro evolution was carried out to enable the hc ligase to operate on a separate template–substrate complex,[74] as is the case for RNA polymerase proteins. The

Table 1: In vitro evolved ribozymes that catalyze formation of a 3′,5′-phosphodiester linkage between two template-bound RNA molecules.

Ligase	Variant	Selection rounds	Length [nt]	k_{cat} [min^{-1}]	[Mg^{2+}] [mM]	pH	*T* [°C]	Ref.
class I	b1-207	14	119	14.4	60	7.4	22	[69, 70]
hc	hc16	10	337	0.26	50	7.5	50	[73]
L1	R8-9	13	96	0.37	60	7.7	25	[77, 80]
R3	R3C	16	71	0.32	25	8.5	23	[78]
DSL	*cis*-DSL-1S	9	140	0.12	50	7.7	37	[79]
P4-P6	clone I	11	153	0.003	80	8.5	37	[83]

template–substrate complex was connected to the ribozyme through a long oligouridylate tether to provide a pseudo-intermolecular reaction format. Eighteen rounds of in vitro evolution were carried out that culminated in a ribozyme that can operate on a completely separate template–substrate complex with a k_{cat} of 0.074 min^{-1} and Michaelis–Menten constant (K_m) of 3.9 μM (measured under the same conditions as above). Both the class I ligase, which operates on an internal template, and the hc ligase, which operates on an external template, have a limited ability to catalyze the addition of NTPs to the 3′-end of an oligonucleotide primer.[74,75] However, as will be discussed in Section 6.1, the class I ligase has been further evolved to operate on an external template and to add as many as 14 successive NTPs.[76]

Three other in vitro evolved ribozymes with 3′,5′-ligase activity are the L1, R3C, and DSL ligases.[77–79] All have a similar architecture based on a central three-helix junction that is offset by several base pairs from the site of ligation, and all operate with a catalytic rate of 0.1–0.4 min^{-1}. The L1 ligase has the unusual property that its activity is dependent on the presence of the DNA primer that was used to initiate the reverse transcription step of selective amplification.[77,80] This is because the primer was present during the ligation reaction and the evolved ligase came to utilize the primer to form a stem structure that is part of the catalytic motif.

Recently the X-ray crystal structure of a modified form of the L1 ligase was solved at 2.6-Å resolution.[81] This is the only ligase ribozyme for which an atomic-resolution structure is available. The structure reveals tertiary contacts between phylogenetically conserved residues near the three-helix junction and nucleotides immediately adjacent to the site of ligation. These contacts involve a GAU-base triple and ionic interactions between a bound Mg^{2+} ion and three phosphate groups. The structure indicates that the ribozyme would not be generalizable to other sequences at the ligation junction and would not be able to bind a separate template–substrate complex. Thus it appears unlikely that the L1 ligase could be evolved to operate as a polymerase ribozyme.

The R3C ligase also does not appear to offer a good starting point for the development of an RNA polymerase ribozyme because it too appears incapable of binding a separate template–substrate complex. However, the R3C ligase has provided an opportunity to explore the relationship between constitutional complexity and the evolvability of catalytic function.[78] The initial version of this ligase was obtained following ten rounds of in vitro evolution, starting with a population of random-sequence RNA molecules that contained adenosine, guanosine, and uridine, but no cytidine. The initial ribozyme was termed the R3 (random 3-letter) ligase. It contains 74 nucleotides (25 A, 32 U, and 17 G) and exhibits a k_{cat} of 0.013 min^{-1} and K_m of 6.2 μM (measured in the presence of 25 mM of $MgCl_2$ and 50 mM of KCl at pH 8.5 and 23 °C). The R3 ligase subsequently was doped with cytidine residues at a frequency of 1% per nucleotide position, subjected to hypermutagenic PCR, and made to undergo 6 additional rounds of in vitro evolution. This resulted in the R3C ligase, which contains 73 nucleotides (22 A, 24 U, 20 G, and 7 C) and has a k_{cat} of 0.32 min^{-1} and K_m of 0.4 μM (measured under the same conditions as above). The added cytidine residues resulted in a 25-fold improvement in catalytic rate, which was attributed to remodeling of the three-helix junction enabled by more stable GC pairs surrounding the junction.

Moving in the opposite direction, the R3 ligase was evolved into a form that contains only two different nucleotide subunits, that is, composed entirely of 2,6-diaminopurine (D) and uracil residues.[82] First all the adenines were replaced by diaminopurine, the GU wobble pairs were replaced by DU pairs, and the unpaired guanines were replaced by an equimolar mixture of diaminopurine and uracil. Then ten rounds of in vitro evolution were carried out, which culminated in the R2 ligase. This ribozyme contains 83 nucleotides (50 D and 33 U) and has a catalytic rate of only 0.0011 min^{-1}. Because of its intense self-complementarity, the ligase is highly susceptible to adopting alternative conformations, with only 6–8% of the molecules folding into an active conformation. The catalytic rate of the R2 ligase is 12-fold slower than that of the R3 ligase, but this still represents a catalytic rate enhancement of 3.6×10^4 compared with the uncatalyzed rate of reaction. The R2 ligase demonstrates how Darwinian evolution can derive macromolecular function even from a highly restricted set of chemical building blocks.

The most recently reported in vitro evolved ribozyme with 3′,5′-ligase activity is the DSL ligase (designed and selected ligase),[79] which was obtained by Inoue and co-workers using an approach similar to that used to obtain the hc ligase. The starting population of RNA molecules contained 30 random-sequence nucleotides that were attached to a structural scaffold derived from a naturally-occurring ribozyme. In addition, tertiary recognition elements were engineering into the molecules that allowed noncovalent interaction between the structural scaffold and the template–substrate complex. During in vitro evolution the template was connected to the 5′-end of the scaffold domain, but it was hoped that the tertiary recognition elements ultimately would allow the ribozyme to operate on a separate template–substrate complex. A total of 10 rounds of in vitro evolution were carried out, which culminated in the DSL ligase. The most active form of the molecule contains 140 nucleotides and has a catalytic rate of 0.12 min^{-1} (measured in the presence of 50 mM of $MgCl_2$ and 200 mM of KCl at pH 7.7 and 37 °C). The ligase can be trimmed to as few as 60 nucleotides and indeed has the ability to react with a separate template–substrate complex, so long as that complex contains the appropriate tertiary recognition elements.

Inoue and co-workers obtained another 3′,5′-ligase ribozyme at about the same time as the DSL ligase by employing a similar approach, but built on a larger structural scaffold and with a less modular design.[83] That ribozyme, which was derived from the P4–P6 domain of the naturally occurring *Tetrahymena* ribozyme, has a catalytic rate of only 0.0027 min^{-1} (measured in the presence of 80 mM of $MgCl_2$ and 50 mM of KCl at pH 8.5 and 37 °C). It has not been subject to further evolutionary optimization and is largely overshadowed by the DSL ligase. The DSL ligase has the potential to be evolved to function as an RNA polymerase. Like the class I and hc ligases, it is able to catalyze the templated

addition of NTPs onto the 3′-end of an oligonucleotide primer.[79] If the DSL ligase ever were to function as a polymerase, however, the tertiary recognition elements, which are fixed in position relative to the template–substrate complex, would need to be replaced by some other recognition element that can accommodate polymerase extension products of increasing length.

6. Evolution of an RNA Polymerase Ribozyme

6.1. From Ligase to Polymerase

Thus far the only ligase ribozyme that has been evolved to function as an RNA-dependent RNA polymerase is the class I ligase.[76] Bartel and co-workers began with the b1-207 form of this ribozyme (Figure 5b), introduced random mutations throughout the molecule, and attached 76 random-sequence nucleotides to its 3′-end. They deleted the internal template region and instead supplied an external template, and they attached the oligonucleotide primer to the 5′-end of the ribozyme through a 5′,5′-phosphodiester linkage. A population of these RNA constructs were challenged to catalyze the template-directed addition of NTPs to the 3′-end of the primer. They were selected based on acquisition of a chemical tag, either *N*6-biotinyl-adenine or 4-thiouracil, contained within the last NTP to be added. Ribozymes that incorporated *N*6-biotinyl-ATP were selected based on binding to streptavidin, while those that incorporated 4-thio-UTP were selected based on interaction of the thiol with *N*-acryloylaminophenylmercuric acetate contained within a polyacrylamide gel, which retarded the gel mobility of reacted ribozymes. The biotin tag led to some undesirable outcomes, but the 4-thiouracil tag proved effective for the selective enrichment of ribozymes with NTP-adding activity.

After ten rounds of in vitro evolution, individuals were cloned from the population and sequenced. A single clone was identified that could add multiple NTPs to the 3′-end of a separate oligonucleotide primer in a template-directed manner.[76] That particular clone was randomly mutagenized and eight additional rounds of in vitro evolution were carried out. Again individuals were isolated from the population and a clone was identified that had especially robust polymerase activity. It subsequently was modified based on comparison with the sequences of the other active clones, thus resulting in the "18.12.23" ribozyme (also referred to as the class-I-derived polymerase). This remarkable ribozyme, which contains 189 nucleotides, has the ability to polymerize as many as 14 successive NTPs on an external RNA template (Figure 6), it exhibits high fidelity of nucleotide addition, and can operate on a variety of different template sequences.

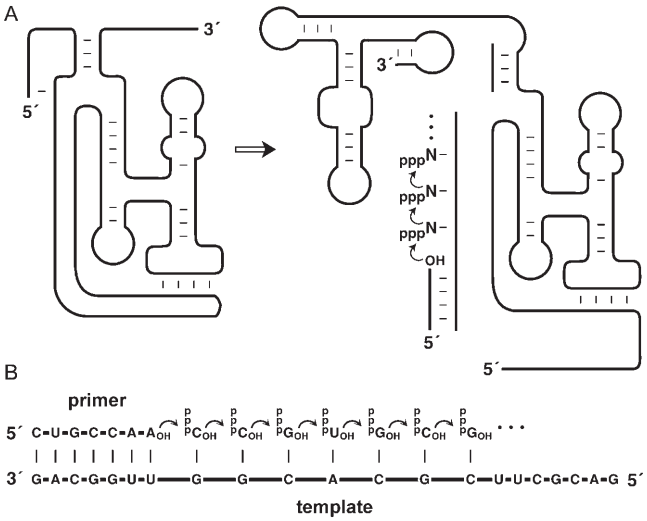


Figure 6. RNA-catalyzed polymerization of RNA. A) The class I ligase ribozyme (left) was evolved to function as an RNA polymerase (right) that catalyzes the addition of NTPs to the 3′-end of a template-bound oligonucleotide primer. B) When a particular 21-nucleotide template and complementary 7-nucleotide primer are employed, the polymerase catalyzes up to 14 successive NTP additions, which results in the synthesis of a fully double-stranded product.

6.2. Toward a More Robust RNA Polymerase

Closer inspection of the polymerase ribozyme reveals some of its key limitations. First, although the k_{cat} for NTP addition is greater than 1 min^{-1}, the K_m for the separate template–primer complex is immeasurably high and in excess of 1 mM.[76,84] In practical terms this means that if one employs typical concentrations of 1 μM template–primer, about 2 hours are required for a productive substrate-binding event. For most templates the ribozyme has little processivity, that is, little ability to add multiple NTPs before dissociating from the template–primer complex.[84] Thus another 2 hours are required for the next productive binding event. The ribozyme is susceptible to hydrolysis of its component phosphodiester linkages, and under the preferred reaction conditions of 200 mM of $MgCl_2$ at pH 8.5 and 22 °C, suffers nonspecific cleavage of one of its phosphodiesters at a rate of about 10^{-2} min^{-1}.[85] Thus in the race between NTP addition and degradation of the ribozyme, it is possible to achieve about 12 NTP additions in 24 h, but not many more, because by then the ribozyme is largely degraded.

A second limitation of the class-I-derived polymerase ribozyme is that, although for one special template it can add up to 14 successive NTPs, for more typical templates it adds only a few NTPs.[84] Even a very subtle change in the sequence of the preferred template dramatically reduces the extent of NTP addition. A third limitation is that, although the fidelity of template copying is high when measured for the full-length products, the overall fidelity is considerably lower because incorporation of the wrong NTP reduces the rate of subsequent extension.[86] A fourth limitation is that the affinity of NTP binding to the template is determined largely by the strength of Watson–Crick pairing, thus requiring high concentrations of NTPs and providing an inherent advantage for GC pairs.[76]

In principle each of these limitations could be overcome with further evolution, and some attempts have been made to do so. Lawrence and Bartel went back to the population of ribozymes obtained after the first three rounds of in vitro

evolution for polymerase activity and set out on a different path.[87] They did not select based on the incorporation of *N*6-biotinyl-adenine residues, which had proven problematic previously, and instead selected based on the incorporation of two successive 4-thiouracil residues through trapping of reacted molecules at the interface between a standard polyacrylamide gel and a mercury-containing polyacrylamide gel. After a total of 11 rounds of evolution, several new ribozyme motifs were identified. A few of these were reasonably good polymerases, but none were as efficient as the previously identified class-I-derived polymerase.

More recently, Zaher and Unrau used a water-in-oil emulsion technique to select ribozyme polymerases based on their activity within isolated compartments.[86] They began with the class-I-derived polymerase and introduced random mutations at a frequency of 3% per nucleotide position within the original ligase domain and 10% per nucleotide position within the added 3′-terminal domain. A population of ribozyme-encoding DNA molecules were distributed among individual compartments that also contained T7 RNA polymerase, the four NTPs, and a template for RNA-catalyzed primer extension. The RNA primer was attached to the 3′-end of the ribozyme-encoding DNA. Transcription occurred within the compartments and the resulting RNA molecules were selected for their ability to extend the primer, thus tagging the DNA molecules that encoded ribozymes responsible for the extension reaction. The tagged DNA molecules were harvested and amplified, and the resulting DNA copies were distributed among a new set of compartments. To explore a large and diverse population of ribozymes, methods were developed for producing 3 L of the water-in-oil emulsion that contained around 10^{15} different individuals.

Following six rounds of selective amplification within compartments, Zaher and Unrau identified a particular ribozyme, designated B6.61, that is around three-fold faster than the previously-identified class-I-derived polymerase.[86] The B6.61 ribozyme still has a poor ability to bind the template–primer complex, but is able to polymerize 20 successive NTPs on an external RNA template. This improvement in the maximum number of added NTPs is primarily due to an increased fidelity of reaction, which reduces the probability of chain termination as a result of the incorporation of the wrong NTP.

In summary, although the development of an RNA polymerase ribozyme that can copy long RNA templates of almost any sequence appears feasible, no such catalyst is yet available. Perhaps it will be achieved through further evolution of descendants of the class I ligase. This would require substantial improvement of the ability of the polymerase to recognize a separate template–substrate complex, and to operate with even greater fidelity and sequence generality. Perhaps an efficient polymerase ribozyme will be obtained through further evolution of the DSL ligase or one of the other ligase ribozymes discussed above. More likely, however, this development will require starting again from a large pool of random-sequence RNA molecules and carrying out extensive in vitro evolution aimed toward a polymerase ribozyme. Thus far no one has evolved a ligase that from the outset was selected for its ability to catalyze NTP addition on a separate template–substrate complex. With ongoing advances in in vitro evolution methods, including the use of water-in-oil emulsions and the exploration of small-molecule cofactors that can assist in RNA catalysis, the development of a robust RNA polymerase ribozyme may yet be achievable.

7. Evolution of Ribozymes by Serial Transfer

If a sufficiently powerful RNA polymerase ribozyme could be developed, then it might be used in place of Qβ replicase to carry out the in vitro evolution of RNA by a serial transfer procedure. Like Qβ replicase, the replicase ribozyme would need to copy both the plus and minus strands of the evolving RNA molecules. This would require some means of separating the two strands of the RNA duplex, for example as Qβ replicase does by causing the newly synthesized strand to fold upon itself rather than remain bound to the template.[88] Unlike Qβ replicase, however, a replicase ribozyme might be able to copy a broad range of template sequences, thus allowing one to evolve RNA molecules with catalytic function, including polymerase-like activity.

7.1. Evolution Becomes Continuous

Most in vitro evolution experiments with ribozymes have been carried out in a stepwise fashion, with extensive manipulation of the reaction materials during successive rounds of selective amplification and mutation. For certain ligase ribozymes, however, it is possible to evolve the RNA molecules in a continuous manner using a serial transfer procedure similar to that employed by Spiegelman (Figure 7).[89] The reaction mixture for the continuous evolution of ribozymes is more complicated than that used to evolve Qβ RNA, requiring input ribozymes, the substrate for RNA-catalyzed ligation, the two polymerase proteins (reverse transcriptase and T7 RNA polymerase) needed for isothermal amplification of RNA, the DNA primer used to

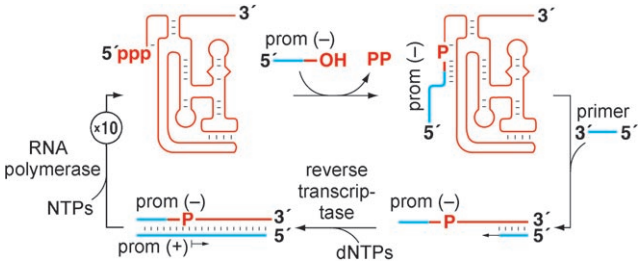


Figure 7. Continuous in vitro evolution of the class I ligase ribozyme.[89] The ribozyme catalyzes the joining of a chimeric DNA–RNA substrate to its 5′ end (DNA: blue, RNA: red). The substrate has the sequence of the T7 RNA polymerase promoter (prom). Reverse transcription of reacted, but not unreacted, ribozymes generates double-stranded molecules that are transcribed by the T7 RNA polymerase to produce multiple copies of progeny RNA. The 5′-triphosphate is restored during transcription, thus allowing the progeny to perform another reaction.

initiate cDNA synthesis, the four dNTPs, the four NTPs, $MgCl_2$, KCl, and buffer. The substrate has the sequence of the T7 RNA polymerase promoter and is a chimeric molecule that contains five or more deoxyribonucleotides at its 5′-end and one or more ribonucleotides at its 3′-end.[90]

In the continuous-evolution mixture, any ribozyme that ligates the substrate to its 5′-end (through the reaction between the 3′-hydroxy group of the substrate and the 5′-triphosphate of the ribozyme) acquires the promoter sequence at its 5′-end. When these reacted RNA molecules are reverse transcribed, the resulting double-stranded material contains both strands of the T7 promoter. Unreacted RNA molecules also are reverse transcribed, thus sequestering the RNA as an RNA–DNA heteroduplex that can no longer react and that lacks a promoter. Material that contains the double-stranded promoter is recognized by T7 RNA polymerase, which generates multiple copies of RNA, each bearing a 5′-triphosphate. The start site for transcription corresponds to the ligation junction for the RNA-catalyzed reaction. Thus the newly-synthesized RNA molecules are returned to a form that is capable of performing another ligation reaction.

All of the events of continuous evolution occur repeatedly within a common reaction mixture until the supply of substrate molecules (or some other limiting reagent) is exhausted. One can then take an aliquot of the reaction mixture and transfer it to a new reaction vessel that contains a fresh supply of materials. Only those ribozymes that are carried over in the aliquot have an opportunity to react in the next mixture. Thus for a particular ribozyme to survive, it must react faster than it is reverse transcribed, it must generate enough copies to compensate for the dilution factor of the transfer, and it must compete with the other ribozymes in the mixture to generate as many copies as possible prior to transfer.

The class I ligase was the first ribozyme that was made to undergo continuous in vitro evolution.[89] The b1-207 form of the class I ligase was modified so that its internal template region would be complementary to a substrate that has the sequence of the T7 promoter. This modified ribozyme reacted poorly with an all-RNA substrate and had no detectable activity with the requisite chimeric DNA–RNA substrate. Fifteen rounds of stepwise evolution were then carried out to restore a catalytic rate of about 0.1 min^{-1} under reaction conditions that are favorable for the class I ligase, but the resulting ribozymes still were substantially less active than the parental b1-207 form under conditions required by reverse transcriptase and T7 RNA polymerase (for example, 25 mM of $MgCl_2$ and 50 mM of KCl at pH 8.5 and 37 °C). Next a rapid evolution procedure was used to adapt the ribozymes to the desired conditions. This procedure alternates between pairs of reaction mixtures, one for RNA catalysis and another for selective isothermal amplification, while progressively narrowing the differences in conditions between the two mixtures. After 100 rounds of this procedure, the population of ribozymes could react with the chimeric substrate under the desired conditions at a rate of greater than 1 min^{-1}.

The molecules that resulted from the rapid evolution procedure were used to initiate continuous in vitro evolution.[89] During the first 24 transfers the incubation time was 1 h and the dilution factor was 1000-fold. As in the experiment of Spiegelman in 1967,[2] the dilution factor was kept constant with successive transfers and the incubation time was decreased as tolerated by the evolving population, eventually reaching 15 min. A total of 100 transfers were carried out, which corresponds to an overall dilution of around 10^{300}-fold (Figure 8). This is a mind-boggling number for a chemical process, but it should be remembered that the system exhibits exponential growth, which is maintained through repeated dilution. Each ligation event results in the production of about 10 copies of the reacted RNA. Thus an overall amplification of 10^{300}-fold corresponds to approximately 300 successive rounds of ligation and selective amplification.

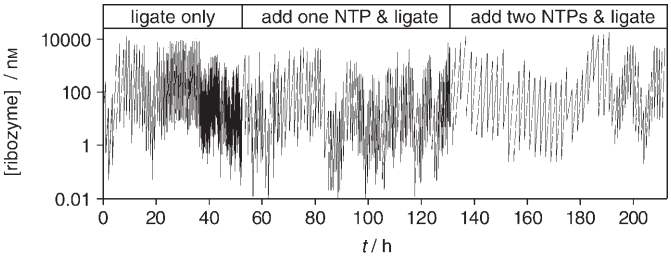


Figure 8. Continuous evolution of ribozymes that catalyze RNA ligation,[89] and subsequently catalyze one or two NTP additions, followed by RNA ligation.[97] A total of 265 transfers were carried out, with a 1000-fold dilution per transfer and a variable amount of time between transfers. The graph shows the concentration of ribozymes before and after each transfer. The overall dilution is calculated as the product of the dilution performed at each transfer.

The ribozymes that emerged after 100 transfers of continuous evolution typically contained about 30 mutations relative to the b1-207 ligase, and about 15 mutations relative to the molecules that were present at the start of continuous evolution.[89] These mutations resulted in about a 10^4-fold increase in the catalytic efficiency of the ribozyme, as a result of a 10^2-fold improvement in k_{cat} and 10^2-fold improvement in K_m. The former was difficult to attribute to any particular mutation, while the latter was largely due to an increase in the length of the internal template region of the ribozyme. One especially well-studied variant that was obtained at the end of the 100 transfers, termed the E100-3 ligase, has a k_{cat} of 21 min^{-1} and K_m of 1.7 μM.[91] It has a doubling time of 1.5 min in the continuous in vitro evolution mixture, thus producing 10^{12} copies per hour under exponential growth conditions.

7.2. New Challenges for Continuously Evolving Ribozymes

Several continuous evolution experiments have been carried out starting with randomized variants of the E100-3 ligase. These have led to evolutionary descendants with altered phenotypes, for example, the ability to operate in the presence of reduced concentrations of $MgCl_2$[92] or to function under acidic (pH 5.8) or alkaline (pH 9.8) conditions.[93] Other studies have used continuous evolution to address classical questions in evolutionary biology, such as the likelihood of

recurrent evolutionary outcomes[94] or the effect of the progressive accumulation of deleterious mutations.[95]

One continuous evolution study,[91] reminiscent of the experiments in which Qβ RNA was evolved to be resistant to EtBr,[14,15] employed a ribozyme-inactivating DNA enzyme[96] to challenge the ribozyme to develop resistance. The DNA enzyme was designed to cleave the ligase ribozyme adjacent to the site of ligation, thus preventing the ribozyme from carrying out its function. The DNA enzyme bound to the ribozyme through Watson–Crick pairing to nucleotides that encompass the internal template region of the ribozyme. These nucleotides cannot be mutated because they are necessary for the ribozyme to recognize its own substrate. The solution to this difficult evolutionary challenge was twofold.[91] First, blocking mutations were acquired at either end of the internal template region (together with compensatory mutations to maintain the proper folded structure of the ribozyme), which reduced the ability of the DNA enzyme to cleave the ribozyme by about 2000-fold. Second, mutations were acquired that improved the K_m of the ribozyme for its own substrate. This allowed the ribozyme to be more completely saturated by its substrate, thus reducing binding by the DNA enzyme, which occurs in competition with binding of the substrate by the ribozyme.

Relevant to the previous discussion of polymerase ribozymes, continuous in vitro evolution has been used to obtain variants of the E100-3 ligase that perform one or two NTP additions followed by ligation.[97] To achieve this, the oligonucleotide substrate was shortened by one or two nucleotides and the ribozyme was required to complete the promoter sequence by the addition of the appropriate NTPs, and then ligate itself to the extended substrate. The ribozymes first were evolved for the ability to perform one NTP addition followed by ligation. This required 5 rounds of stepwise evolution followed by 105 transfers of continuous evolution with a 1000-fold dilution per transfer (Figure 8). Then the ribozymes were evolved for the ability to carry out two NTP additions followed by ligation, which required 8 rounds of stepwise evolution and 60 transfers of continuous evolution with a 1000-fold dilution per transfer. A typical ribozyme that was isolated from the final evolved population, designated E278-19, contained 30 mutations relative to the E100-3 ligase. It performs two template-directed NTP additions followed by ligation at an overall rate of 3×10^{-4} min^{-1} (measured in the presence of 2 mM of each NTP, 25 mM of $MgCl_2$, and 50 mM of KCl at pH 8.5 and 37 °C).

Attempts to use continuous evolution to develop ribozymes that perform three or more NTP additions followed by ligation proved unsuccessful.[97] When the substrate was shortened by three nucleotides, the evolving molecules were allowed to escape the selection constraint by generating nonstandard promoter sequences that still were functional in transcription. Thus it may be difficult to use continuous evolution to develop efficient RNA polymerase ribozymes. Continuous evolution also is limited in the range of chemical reactions that it can support. In principle, continuous evolution could be applied to other bond-forming reactions that result in attachment of a promoter-containing substrate to the 5′-end of the ribozyme. However, the newly-formed bond must be traversable by reverse transcriptase and the reactive group at the 5′-end of the ribozyme must be restored during transcription.

At present, in vitro evolution by serial transfer has only been demonstrated for Qβ RNA, certain nonfunctional nucleic acids that can be amplified isothermally,[98,99] and fast-reacting ligase ribozymes. Serial transfer is widely employed as a method for propagating bacterial and eukaryotic cells, and has long been practiced in the breeding of plants and domesticated animals. However, those are biological processes that, by definition, cannot be performed in vitro.

8. In Vitro Evolution Today and Tomorrow

It has been difficult to establish a second ribozyme, unrelated to the class I ligase, that is capable of undergoing continuous in vitro evolution. No other ligase ribozyme has a catalytic rate nearly as fast as that of the class I ligase. A fast catalytic rate is necessary for continuous evolution because the ribozyme must react before it becomes inactivated by reverse transcription.

Recently a second lineage of continuously evolving ribozymes was established based on optimized variants of the DSL ligase.[100] This effort required configuring the molecule so that it could react with a substrate that has the sequence of the T7 RNA polymerase promoter, but this resulted in a loss of activity. Accordingly, stepwise evolution was carried out to reestablish a low level of activity. Then a random-sequence domain of 35 nucleotides was added, and the catalytic rate was optimized by selecting molecules that could react in times as short as 15 milliseconds. These very short reaction times were achieved using a quench-flow apparatus, which resulted in a population of ribozymes that was capable of initiating continuous in vitro evolution and was carried through 80 transfers, with an overall dilution of more than 10^{200}-fold.[100]

With two distinct "species" of continuously evolving ribozymes available, it now is possible to conduct in vitro coevolution studies in which the two ribozymes are made to operate within a common environment. This will provide an opportunity for exploring the possibility of competition and cooperation among functional molecular species. In vitro studies in "molecular ecology" have been carried out previously using two different nonfunctional nucleic acids that undergo isothermal amplification in the same mixture.[101,102] In one study, the cDNA form of one molecule served as the cDNA primer for the other, resulting in a coupled system that could operate briefly before succumbing to the emergence of parasitic molecules.[101] In another study, the cDNA form of two different molecules served as primers for each other, which resulted in a cooperatively coupled system that also quickly succumbed to parasites.[102] Long-term coupled behavior in an in vitro evolution system will likely require each species to perform its own function in order to be eligible for amplification, thus reducing susceptibility to nonfunctional parasites.

Another recent advance pertaining to continuous in vitro evolution is the development of a microfluidic device for

carrying out automated serial dilutions.[103] This device has been used to conduct the continuous evolution of ligase ribozymes within sub-microliter volumes that are confined to fluidic circuits within a fabricated glass wafer.[104] The evolving population is monitored continuously using a confocal laser microscope and the intercalating dye thiazole orange, which provides a fluorescent measurement of ribozyme concentration. When a chosen threshold concentration is reached, computer-controlled microvalves are activated to isolate a fixed aliquot of the reaction mixture and then dilute it with a mixture containing a fresh supply of reagents. This process of growth and dilution can be repeated many times, analogous to serial transfer performed by manual methods. Evolution on a chip, however, is far more precise and reproducible than manual techniques.[103] The microfluidic evolution device can maintain a population of 10^9–10^{11} individuals through an exponential growth phase of variable length, then perform a 5- to 25-fold dilution in 20 s before resuming exponential growth.

A robotic workstation has been used to carry out the in vitro selection of RNA aptamers.[105] More recently, a prototype microfluidic device was reported that automates the steps of aptamer selection, although no aptamer has yet been generated using this device.[106] As discussed in Section 3.1, the in vitro evolution of Qβ RNA has been conducted using a robotic serial-transfer apparatus.[16] It is likely that Qβ RNA evolution could be carried out in a microfluidic device similar to that used to evolve ligase ribozymes. Ultimately microfluidic devices will be applied to polymerase and ligase ribozymes, thus offering the possibility of conducting many thousands of successive rounds of in vitro evolution without operator intervention. Once the design of a microfluidic chip has been validated, individual chips can be produced at low cost, and because each chip contains 24 separate serial dilution circuits, it will be possible to conduct parallel and highly longitudinal studies with many consecutive steps of evolving ribozymes.

An overarching goal of in vitro RNA evolution research is to develop a system that can evolve in a self-sustained manner, without the aid of protein polymerases or any other informational macromolecule that is not part of the evolving system. By some definitions this would be the realization of life in the laboratory.[107–109] A promising route toward achieving this goal appears to be the development of a robust RNA polymerase ribozyme that can function as an RNA replicase. As discussed above, this likely will require substantial in vitro evolution that begins with either an existing ligase ribozyme or a population of random-sequence RNA molecules. Fortunately, in vitro evolution methods are becoming increasingly powerful, raising expectations that RNA molecules with the requisite catalytic properties can be attained.

Self-sustained evolution of a replicase ribozyme requires the catalytic strand to copy both plus and minus strand RNA molecules that correspond to the catalyst and its complement, respectively. It has been suggested how a ribozyme might literally copy itself,[110] but more likely the ribozyme would copy other molecules that have the same or complementary sequence as itself. Once self-sustained RNA evolution begins to operate, the ribozymes would enjoy a strong selective advantage if they specifically recognized and amplified only those RNA molecules that are closely related to themselves. One suggestion for how this might occur is through recognition of a tag sequence, perhaps resembling a tRNA, located at the end of both the plus and minus strands.[111] Specific recognition by the replicase, whether it is directly related to substrate binding or a consequence of confining the enzyme and substrate to a common locale, would result in the expression of "selective preference", as first described by Spiegelman and co-workers for the replication of RNA viruses.[3] This property would become engrained through further in vitro evolution, thus causing the replicase to become more restrictive with regard substrate sequence. Thus a system for the artificial selection of molecules, in which the experimenter determines the selection constraints, would eventually give way to a form of natural selection in the laboratory, in which functional molecules evolve on their own and without direction toward a particular phenotypic trait.

I am grateful to my co-workers who have shared in the development and application of in vitro RNA evolution. I am also grateful to Leslie Orgel and Richard Lerner for their many helpful suggestions over the years and for their comments on this Review. Our work has been supported by the Exobiology Program of the National Aeronautics and Space Administration and by the National Institutes of Health.

Received: March 29, 2007
Published online: July 16, 2007

[1] C. R. Darwin, *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, John Murray, London, **1859**.
[2] D. R. Mills, R. L. Peterson, S. Spiegelman, *Proc. Natl. Acad. Sci. USA* **1967**, *58*, 217–224.
[3] I. Haruna, K. Nozu, Y. Ohtaka, S. Spiegelman, *Proc. Natl. Acad. Sci. USA* **1963**, *50*, 905–911.
[4] I. Haruna, S. Spiegelman, *Proc. Natl. Acad. Sci. USA* **1965**, *54*, 579–587.
[5] I. Haruna, S. Spiegelman, *Science* **1965**, *150*, 884–886.
[6] S. Spiegelman, I. Haruna, I. B. Holland, G. Beaudreau, D. Mills, *Proc. Natl. Acad. Sci. USA* **1965**, *54*, 919–927.
[7] F. Meyer, H. Weber, C. Weissmann, *J. Mol. Biol.* **1981**, *153*, 631–660.
[8] T. Nishihara, D. R. Mills, F. R. Kramer, *J. Biochem.* **1983**, *19*, 228–236.
[9] D. Brown, L. Gold, *Biochemistry* **1995**, *34*, 14765–14774.
[10] C. Tuerk, L. Gold, *Science* **1990**, *249*, 505–510.
[11] G. Klein, *The Atheist and the Holy City: Encounters and Reflections* (translated by T. Friedmann, I. Friedmann), MIT Press, Cambridge, MA, **1990**, pp. 37–52.
[12] F. Wöhler, *Ann. Chim. Phys.* **1828**, *37*, 330–333.
[13] R. Levisohn, S. Spiegelman, *Proc. Natl. Acad. Sci. USA* **1969**, *63*, 805–811.
[14] R. Saffhill, H. Schneider-Bernloehr, L. E. Orgel, S. Spiegelman, *J. Mol. Biol.* **1970**, *51*, 531–539.
[15] F. R. Kramer, D. R. Mills, P. E. Cole, T. Nishihara, S. Spiegelman, *J. Mol. Biol.* **1974**, *89*, 719–736.
[16] G. Strunk, T. Ederhof, *Biophys. Chem.* **1997**, *66*, 193–202.
[17] C. K. Biebricher, M. Eigen, W. C. Gardiner, *Biochemistry* **1985**, *24*, 6550–6560.

[18] H. V. Chetverina, A. B. Chetverin, *Nucleic Acids Res.* **1993**, *21*, 2349 – 2353.
[19] G. J. Bauer, J. S. McCaskill, H. Otten, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 7937 – 7941.
[20] A. B. Chetverin, H. V. Chetverina, A. A. Demidenko, V. I. Ugarov, *Cell* **1997**, *88*, 503 – 513.
[21] J. S. McCaskill, G. J. Bauer, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 4191 – 4195.
[22] E. A. Miele, D. R. Mills, F. R. Kramer, *J. Mol. Biol.* **1983**, *171*, 281 – 295.
[23] P. M. Lizardi, C. E. Guerra, H. Lomeli, I. Tussie-Luna, F. R. Kramer, *Biotechnology* **1988**, *6*, 1197 – 1202.
[24] Y. Wu, D. Y. Zhang, F. R. Kramer, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 11769 – 11773.
[25] R. K. Saiki, S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich, N. Arnheim, *Science* **1985**, *230*, 1350 – 1354.
[26] R. K. Saiki, D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, H. A. Erlich, *Science* **1988**, *239*, 487 – 491.
[27] J. C. Guatelli, K. M. Whitfield, D. Y. Kwoh, K. J. Barringer, D. D. Richman, T. R. Gingeras, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 1874 – 1878.
[28] J. Compton, *Nature* **1991**, *350*, 91 – 92.
[29] A. D. Ellington, J. W. Szostak, *Nature* **1990**, *346*, 818 – 822.
[30] L. Gold, B. Polisky, O. Uhlenbeck, M. Yarus, *Annu. Rev. Biochem.* **1995**, *64*, 763 – 797.
[31] D. S. Wilson, J. W. Szostak, *Annu. Rev. Biochem.* **1999**, *68*, 611 – 647.
[32] *The Aptamer Handbook: Functional Oligonucleotides and Their Applications* (Ed.: S. Klussmann), Wiley-VCH, Weinheim, **2006**.
[33] L. C. Griffin, G. F. Tidmarsh, L. C. Bock, J. J. Toole, L. L. Leung, *Blood* **1993**, *81*, 3271 – 3276.
[34] J. Floege, T. Ostendorf, U. Janssen, M. Burg, H. H. Radeke, C. Vargeese, S. C. Gill, L. S. Green, N. Janjic, *Am. J. Pathol.* **1999**, *154*, 169 – 179.
[35] C. P. Rusconi, E. Scardino, J. Layzer, G. A. Pitoc, T. L. Ortel, D. Monroe, B. A. Sullenger, *Nature* **2002**, *419*, 90 – 94.
[36] H. A. D. Lagasse, P. G. Merlino, M. C. Gilbert, J. R. Killough, K. Mills, R. M. Boomer, S. D. Lewis III, A. S. Makim, C. R. Benedict, T. G. McCauley, J. B. Rottman, H. N. Marsh, J. L. Diener, *Arterioscler. Thromb. Vasc. Biol.* **2005**, *25*, 66.
[37] P. Kobelt, S. Helmling, A. Stengel, B. Wlotzka, V. Andresen, B. F. Klapp, B. Wiedenmann, S. Klussmann, H. Mönnikes, *Gut* **2006**, *55*, 788 – 792.
[38] J. Ruckman, L. S. Green, J. Beeson, S. Waugh, W. L. Gillette, D. D. Henninger, L. Claesson-Welsh, N. Janji, *J. Biol. Chem.* **1998**, *273*, 20556 – 20567.
[39] E. W. M. Ng, D. T. Shima, P. Calias, E. T. Cunningham Jr., D. R. Guyer, A. P. Adamis, *Nat. Rev. Drug Discovery* **2006**, *5*, 123 – 132.
[40] D. L. Robertson, G. F. Joyce, *Nature* **1990**, *344*, 467 – 468.
[41] M. A. Sirover, L. A. Loeb, *J. Biol. Chem.* **1977**, *252*, 3605 – 3610.
[42] P. Keohavong, W. G. Thilly, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 9253 – 9257.
[43] K. A. Eckert, T. A. Kunkel, *Nucleic Acids Res.* **1990**, *18*, 3739 – 3744.
[44] M. Eigen, *Naturwissenschaften* **1971**, *58*, 465 – 523.
[45] J. McCaskill, *Biol. Cybernetics* **1984**, *50*, 60 – 73.
[46] M. Eigen, J. McCaskill, P. Schuster, *J. Phys. Chem.* **1988**, *92*, 6881 – 6891.
[47] P. Schuster, J. Swetina, *Bull. Math. Biol.* **1988**, *50*, 635 – 660.
[48] D. W. Leung, E. Chen, D. V. Goeddel, *Technique* **1989**, *1*, 11 – 15.
[49] R. C. Cadwell, G. F. Joyce in *PCR Primer: A Laboratory Manual*, 2nd ed. (Eds.: C. W. Dieffenbach, G. S. Dveksler), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, **2003**, pp. 453 – 458.
[50] R. C. Cadwell, G. F. Joyce, *PCR Methods Applic.* **1992**, *2*, 28 – 33.
[51] J.-P. Vartanian, M. Henry, S. Wain-Hobson, *Nucleic Acids Res.* **1996**, *24*, 2627 – 2631.
[52] J. Cline, H. Hogrefe, *Strategies* **2000**, *13*, 157 – 162.
[53] B. D. Biles, B. A. Connolly, *Nucleic Acids Res.* **2004**, *32*, e176.
[54] J. D. Hermes, S. M. Parekh, S. C. Blacklow, H. Köster, J. R. Knowles, *Gene* **1989**, *84*, 143 – 151.
[55] D. P. Bartel, J. W. Szostak, *Science* **1993**, *261*, 1411 – 1418.
[56] B. Zhang, T. R. Cech, *Nature* **1997**, *390*, 96 – 100.
[57] P. J. Unrau, D. P. Bartel, *Nature* **1998**, *395*, 260 – 263.
[58] T. M. Tarasow, S. L. Tarasow, B. E. Eaton, *Nature* **1997**, *389*, 54 – 57.
[59] B. Seelig, A. Jäschke, *Chem. Biol.* **1999**, *6*, 167 – 176.
[60] G. Sengle, A. Eisenführ, P. S. Arora, J. S. Nowick, M. Famulok, *Chem. Biol.* **2001**, *8*, 459 – 473.
[61] S. Fusz, A. Eisenführ, S. G. Srivatsan, A. Heckel, M. Famulok, *Chem. Biol.* **2005**, *12*, 941 – 950.
[62] A. A. Beaudry, G. F. Joyce, *Science* **1992**, *257*, 635 – 641.
[63] T. Pan, O. C. Uhlenbeck, *Biochemistry* **1992**, *31*, 3887 – 3895.
[64] S.-M. Chun, S. Jeong, J.-M. Kim, B.-O. Chong, Y.-K. Park, H. Park, J. Yu, *J. Am. Chem. Soc.* **1999**, *121*, 10844 – 10845.
[65] P. A. Lohse, J. W. Szostak, *Nature* **1996**, *381*, 442 – 444.
[66] D. Saran, D. M. Held, D. H. Burke, *Nucleic Acids Res.* **2006**, *34*, 3201 – 3208.
[67] R. L. Coppins, S. K. Silverman, *J. Am. Chem. Soc.* **2004**, *126*, 16426 – 16432.
[68] W. E. Purtha, R. L. Coppins, M. K. Smalley, S. K. Silverman, *J. Am. Chem. Soc.* **2005**, *127*, 13124 – 13125.
[69] E. H. Ekland, J. W. Szostak, D. P. Bartel, *Science* **1995**, *269*, 364 – 370.
[70] E. H. Ekland, D. P. Bartel, *Nucleic Acids Res.* **1995**, *23*, 3231 – 3238.
[71] N. H. Bergman, W. K. Johnston, D. P. Bartel, *Biochemistry* **2000**, *39*, 3115 – 3123.
[72] R. Rohatgi, D. P. Bartel, J. W. Szostak, *J. Am. Chem. Soc.* **1996**, *118*, 3332 – 3339.
[73] L. Jaeger, M. C. Wright, G. F. Joyce, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 14712 – 14717.
[74] K. E. McGinness, G. F. Joyce, *Chem. Biol.* **2002**, *9*, 297 – 307.
[75] E. H. Ekland, D. P. Bartel, *Nature* **1996**, *382*, 373 – 376.
[76] W. K. Johnston, P. J. Unrau, M. S. Lawrence, M. E. Glasner, D. P. Bartel, *Science* **2001**, *292*, 1319 – 1325.
[77] M. P. Robertson, A. D. Ellington, *Nat. Biotechnol.* **1999**, *17*, 62 – 66.
[78] J. Rogers, G. F. Joyce, *RNA* **2001**, *7*, 395 – 404.
[79] Y. Ikawa, K. Tsuda, S. Matsumura, T. Inoue, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13750 – 13755.
[80] M. P. Robertson, J. R. Hesselberth, A. D. Ellington, *RNA* **2001**, *7*, 513 – 523.
[81] M. P. Robertson, W. G. Scott, *Science* **2007**, *315*, 1549 – 1553.
[82] J. S. Reader, G. F. Joyce, *Nature* **2002**, *420*, 841 – 844.
[83] W. Yoshioka, Y. Ikawa, L. Jaeger, H. Shiraishi, T. Inoue, *RNA* **2004**, *10*, 1900 – 1906.
[84] M. S. Lawrence, D. P. Bartel, *Biochemistry* **2003**, *42*, 8748 – 8755.
[85] Y. Li, R. R. Breaker, *J. Am. Chem. Soc.* **1999**, *121*, 5364 – 5372.
[86] H. S. Zaher, P. J. Unrau, *RNA* **2007**, *13*, 1017 – 1026.
[87] M. S. Lawrence, D. P. Bartel, *RNA* **2005**, *11*, 1173 – 1180.
[88] C. Dobkin, D. R. Mills, F. R. Kramer, S. Spiegelman, *Biochemistry* **1979**, *18*, 2038 – 2044.
[89] M. C. Wright, G. F. Joyce, *Science* **1997**, *276*, 614 – 617.
[90] K. E. McGinness, G. F. Joyce, *J. Biol. Chem.* **2002**, *277*, 2987 – 2991.
[91] P. Ordoukhanian, G. F. Joyce, *Chem. Biol.* **1999**, *6*, 881 – 889.

[92] T. Schmitt, N. Lehman, *Chem. Biol.* **1999**, *6*, 857 – 869.
[93] H. Kühne, G. F. Joyce, *J. Mol. Evol.* **2003**, *57*, 292 – 298.
[94] N. Lehman, *Artif. Life* **2004**, *10*, 1 – 22.
[95] S. J. Soll, C. D. Arenas, N. Lehman, *Genetics* **2006**, *175*, 267 – 275.
[96] S. W. Santoro, G. F. Joyce, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 4262 – 4266.
[97] K. E. McGinness, M. C. Wright, G. F. Joyce, *Chem. Biol.* **2002**, *9*, 585 – 596.
[98] R. R. Breaker, G. F. Joyce, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6093 – 6097.
[99] F. Oehlenschläger, M. Eigen, *Origins Life Evol. Biosphere* **1997**, *27*, 437 – 457.
[100] S. B. Voytek, G. F. Joyce, unpublished results.
[101] B. Wlotzka, J. S. McCaskill, *Chem. Biol.* **1997**, *4*, 25 – 33.
[102] R. Ehricht, T. Ellinger, J. S. McCaskill, *Eur. J. Biochem.* **1997**, *243*, 358 – 364.
[103] B. M. Paegel, W. H. Grover, A. M. Skelley, R. A. Mathies, G. F. Joyce, *Anal. Chem.* **2006**, *78*, 7522 – 7527.
[104] B. M. Paegel, G. F. Joyce in *Ninth International Conference on Miniaturized Systems for Chemistry and Life Sciences* (Eds.: K. F. Jensen, J. Han, D. J. Harrison, J. Voldman), Transducer Research Foundation, Boston, MA, **2005**, pp. 28 – 30.
[105] J. C. Cox, A. D. Ellington, *Bioorg. Med. Chem.* **2001**, *9*, 2525 – 2531.
[106] G. Hybarger, J. Bynum, R. F. Williams, J. J. Valdes, J. P. Chambers, *Anal. Bioanal. Chem.* **2005**, *384*, 191 – 198.
[107] C. Sagan, *Encyclopaedia Britannica*, *Vol. 22*, Encyclopaedia Britannica, Chicago, **1998**, pp. 964 – 981.
[108] G. F. Joyce in *Origins of life: the central concepts* (Eds.: D. W. Deamer, G. R. Fleischaker), Jones and Bartlett, Boston, MA, **1994**, pp. xi–xii.
[109] C. F. Chyba, G. D. McDonald, *Annu. Rev. Earth Planet. Sci.* **1995**, *23*, 215 – 249.
[110] N. R. Pace, T. L. Marsh, *Origins Life Evol. Biosphere* **1985**, *16*, 97 – 116.
[111] A. M. Weiner, N. Maizels, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 7383 – 7387.
[112] C. K. Biebricher, R. Luce, *EMBO J.* **1992**, *11*, 5129 – 5135.

