Evolution of Biological Catalysis: Ribozyme to RNP Enzyme

T.R. Cech

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

Correspondence: thomas.cech@colorado.edu

The organizers of this Cold Spring Harbor Symposium made the bold and exciting decision to cover the topic of evolution from a broad perspective, ranging from prebiotic chemistry through the evolution of molecules, cells, organisms, cognitive systems, ecosystems, and societies. To this end, they convened a broad array of scientists who normally attend more specialized conferences and therefore had little previous opportunity to interact. Now, faced with the prospect of contributing a chapter to the Symposium volume, I decided to take a cue from the organizers and write something with minimal technical jargon that aims to be broadly accessible to a wide range of scientists. This approach carries the danger of superficiality, so to counteract this limitation, I have included some key citations to the technical literature. In addition, entire volumes have been written on my topic of the evolution of RNA and RNPs, and the reader is encouraged to explore The RNA World (Gesteland et al. 2006) for detailed descriptions of the topics covered here.

My goal, then, is to consider plausible pathways for the origins of the biocatalysts present in all living beings on Earth—the enzymes that allow cells and organisms to grow, move, take in nutrients, perform metabolism, respond to their environment, and reproduce. I begin by reviewing what we know about contemporary protein enzymes and RNP enzymes, which provides some guidance regarding how they might have evolved. I then summarize some missing information, the lack of which limits our confidence about early events in the evolution of catalysis. Notwithstanding these limitations, I then propose reasonable scenarios for the evolution of an “RNP World,” in which complexes of ribonucleic acid and protein catalyzed the biochemical reactions necessary for life. Finally, I speculate about the future of RNP enzymes.

WHAT WE KNOW ABOUT BIOLOGICAL CATALYSTS

The most detailed knowledge we have concerning biological catalysis of course comes from contemporary organisms. Furthermore, even though this current knowledge is incomplete, we can continue to explore and to experiment to fill gaps in our understanding of today’s biocatalysts, approaches that are much more difficult to apply to prehistoric events and especially inaccessible for events surrounding the origins of life some four billion years ago. What we know is that in the bacterium Escherichia coli, catalysis is performed by ~2000 different protein enzymes and just a few known RNP enzymes. Baker’s yeast has more than a dozen known RNP enzymes, yet protein enzymes still outnumber RNP enzymes by perhaps 100 to 1.

The numerical inferiority of RNP enzymes, however, belies their importance. For example, consider the four major, essential activities of genetic material (DNA and RNA):

• Replication (reproducing genomic DNA)
• Transcription (copying DNA information into RNA)
• RNA splicing (connecting informational stretches of RNA and deleting interrupting segments) and other posttranscriptional modifications
• Translation (using RNA information to synthesize proteins)

Of these four activities, the first two are performed by protein enzymes (DNA polymerases and RNA polymerases), whereas the latter two rely on RNP enzymes (the spliceosome and the ribosome). Thus, RNP enzymes are not peripheral but are central to contemporary biology.

This mention of the spliceosome and the ribosome provides a good opportunity to define “RNP enzyme” as a catalyst that contains both essential RNA and protein.

components. Whether a component is essential for function can be determined by a genetic knockout experiment. For example, in the case of the spliceosome, five small nuclear RNAs (U1, U2, U4, U5, and U6) are essential for splicing (Guthrie 1991), as are numerous protein subunits including U2AF (U2 RNA-associated factor) (Ruskin et al. 1988) and Prp8 (Pena et al. 2008). In the case of the ribosome, catalysis of peptide bond formation requires the large-subunit rRNA, and several large-subunit ribosomal proteins (notably L2, L15 and L16) are also essential, although none of these is itself a peptidyl transferase (Noller 1993). Note that this definition of the RNP enzyme does not distinguish which macromolecule contains the active site that positions the reacting groups and accelerates the chemical transformation. In the case of the ribosome, the active site for peptidyl transfer is thought to be composed entirely or at least mostly of RNA (Noller 1993; Nissen et al. 2000; Voorhees et al. 2009). The spliceosome is also thought to have an RNA active site (Toor et al. 2008; Valadkhan et al. 2009). The RNP enzyme telomerase, which catalyzes the extension of DNA at the ends of chromosomes, has an active site for the chemical step in the reaction composed of three aspartic acid residues in its TERT (telomerase reverse transcriptase) protein subunit (Lingner et al. 1997), yet catalysis of the overall reaction cycle requires the intimate collaboration between protein and RNA (Greider and Greider and Blackburn 1989; Tzafi et al. 2000; Miller and Collins 2002; Qiao and Cech 2008). Thus, although RNP enzymes require both their RNA and their protein constituents, there is a diverse array of ways in which the two different types of macromolecule collaborate to achieve catalysis.

QUESTIONS REGARDING EARLY EVENTS IN THE EVOLUTION OF BIOLOGICAL CATALYSIS

What was the environment? Any hypothesis about the origin of the first self-replicating system requires assumptions about the environmental conditions: What gases, dissolved molecules, and ions were present, what the temperature, pH, and oxidation-reduction potential might have been, and what sort of radiation flux (visible, UV, and ionizing radiation) was present. More specifically, if RNA is to be considered plausible as a prebiotic self-replicator, the environment must have avoided high pH, high temperature, and high concentrations of certain divalent cations that hydrolyze RNA. Furthermore, oxidative conditions can destroy the bases in RNA, low pH can promote depurination, and UV irradiation results in pyrimidine dimer formation. Unfortunately, there is little certainty about the environmental parameters. For example, the original Miller and Urey (1959) prebiotic simulation experiments started with methane, ammonia, water, and hydrogen and resulted in production of high concentrations of natural amino acids. However, these highly reducing conditions were later considered unrealistic. Recent experiments with nonreducing mixtures including nitrogen and carbon dioxide have shown production of amino acids, especially when oxidation is inhibited (Cleave et al. 2008).

What were the first self-replicating molecules? The criteria for a primordial self-replicator include having monomer units that could reasonably be assembled under prebiotic conditions (Powner et al. 2009), ability to undergo both random and template-dependent polymerization (Inoue and Orgel 1983), and ability to be replicated (A → A + A) with enough fidelity to maintain information content (Eigen 1971). Another key criterion is that there must be a plausible evolutionary pathway from the initial self-replicators to contemporary nucleic acids; i.e., even if one can show that mineral surfaces can self-replicate, it is not clear how they would be replaced by nucleic acids, whereas the principle of complementary base pairing makes it much easier to conceive of one nucleic acid system evolving into another (RNA-RNA → RNA-DNA → DNA-DNA). Thus, the concept of a primordial RNA World, with RNA providing the information and catalyzing its own replication, is attractive. Although it has been argued that RNA is too sophisticated to be a plausible candidate for the first self-replicating informational molecule (Joyce and Orgel 2006), the recent breakthrough in recapitulating pyrimidine nucleotide synthesis by Powner et al. (2009) has generated optimism that RNA might in fact have arisen spontaneously. Alternatively, an RNA World may have been preceded by a simpler self-replicating molecule. Some non-RNA nucleic acids that have been considered as plausible prebiotic predecessors include pyranosyl-RNA (Pitsch et al. 1995), peptide nucleic acid (Egholm et al. 1993), and phosphoramidate-linked glycerol nucleic acid (Chen et al. 2009).

How was chirality established? Any organic molecule containing a carbon atom bonded to four nonidentical chemical groups has a mirror-image form that is distinct from the original molecule, just as one’s right hand is distinct from one’s left hand. Two mirror-image molecules are called enantiomers. If a molecule contains more than one chiral center, there will be multiple stereoisomers called diastereomers. Undirected chemical reactions produce mixtures of these enantiomers or diastereomers, whereas biological reactions are stereochemically pure and need to be stereochemically pure in order to be useful. The requirement for chiral purity arises because if X needs to interact with Y to form X-Y, the enantiomer of X (X’) is unlikely to have any specific interaction with Y because its interacting atoms will be in the wrong special arrangement. Even worse, in a polymeric nucleic acid, the presence of even an occasional monomer unit of the opposite chirality would greatly perturb its structure and function. The discovery of some organic chemical reactions that give rise (randomly) to a single chiral product when subjected to phase transitions may provide clues about the origin of chirality in biochemical molecules (Viedma et al. 2008; Blackmond 2009).

How early were self-replicators encapsulated? It is frequently pointed out that it is difficult to launch natural selection in a complex mixture of freely diffusing molecules; only after self-replicating systems are isolated from their neighbors will their own success accrue to them and give them a selective advantage for survival and reproduction. Modern cells achieve this separation by means of a cell membrane consisting in part of a phospholipid bilayer and in some cases by a cell wall. Primordial self-replicating systems could have been encapsulated as “protocells,”
surrounded by fatty acid membranes (Mansy et al. 2008). Other means of separation, such as being sequestered within a porous matrix or being suspended as aerosol particles (Dobson et al. 2000), have been considered.

PLAUSIBLE SCENARIOS FOR THE EVOLUTION OF AN RNP WORLD

Despite our substantial ignorance regarding the key issues described above, we can still describe plausible pathways for the evolution of an RNP World because we can rely on more recent evolutionary events to provide paradigms. First, it seems reasonable that primitive ribozymes would have encountered and formed complexes with existing small molecules, peptides, and (later) proteins and that natural selection would then occur at the level of the RNP rather than the RNA (Noller 2004; Cech 2009). Second, the ribosome is so similar in all organisms that natural selection would then occur at the level of the RNP as well. These steps are summarized in Figure 1.

The Early RNP World

Laboratory “test tube evolution” experiments are frequently performed with pure RNA, but it is difficult to conceive of a natural system in which RNA monomers and polymers would be synthesized without there also being an enormous collection of other organic molecules, including sugars, amino acids, and short peptides consisting of D and L amino acids (see discussion of chirality above) in branched chains as well as linear arrays. Many of these small molecules bind to RNA; natural examples include aminoglycoside antibiotics binding to rRNA, small metabolites binding to riboswitches, guanosine and arginine binding to group I introns, and arginine and short arginine-rich peptides binding to the human immunodeficiency virus trans-activating response region (HIV TAR) RNA (Yarus 1988; Puglisi et al. 1992).

Some protein enzymes bind cofactors (nicotinamide adenine dinucleotide, coenzyme A, pyridoxyl phosphate) and use them in catalysis. Many of these coenzymes are modified ribonucleotides, perhaps molecular fossils of the early RNA World (White 1976). Just as modern proteins expand their catalytic repertoire using these coenzymes, it seems obvious that RNA could have expanded its catalytic repertoire by using bound organic molecules. Indeed, because binding of small molecules by RNA seems chemically inevitable (see previous paragraph), natural selection must have occurred at the level of RNA–small-molecule complexes. If a prevalent small molecule inhibited an RNA molecule’s function, it would have provided negative selection for that particular RNA, whereas any RNA that bound a small molecule that stabilized its active conformation or extended its functional repertoire would have undergone positive selection.

The Middle RNP World

Evolution of RNP enzymes was constrained when RNA had to make use of whatever amino acids and peptides were present in the environment. Thus, a key event occurred when RNA stumbled upon the ability to synthesize reproducibly those peptides that were useful to its functions. This step is sometimes envisioned as the advent of a very primitive version of today’s ribosome: peptide bond formation being directed by a code provided by a separate mRNA. Although the invention of a primitive ribosome was clearly a breakthrough event, the earliest production of specific peptides may have involved RNAs that always bound and joined the same two activated amino acids. Indeed, the d-Ala-d-Ala ligase that contributes to synthesis of bacterial cell wall peptidoglycans is an example of non-mRNA-directed peptide synthesis. It provides a protein analogy to early RNA enzymes that may have catalyzed peptide ligation.

The Recent RNP World

RNP enzymes have been created more recently, since the advent of the common ancestor to extant life. Once created, they continue to evolve. I describe two examples of RNP evolution, one in which a ribozyme acquired a
protein and another in which a protein enzyme acquired an RNA. Both stories are a bit like historical novels: The details are fictional but plausible, because they are embedded in a context of fact. In each case, the most parsimonious explanation for an observation is given preference over explanations that require multiple unsubstantiated steps.

The first story concerns the evolution of an RNP enzyme from an intron RNA that underwent self-splicing in the absence of protein. The ribozyme is considered the ancestral form, because all of the group I ribozymes share sequence, structure, and mechanistic features, whereas their associated proteins are diverse and largely unrelated. Consider a moderately active group I intron that was able to bind weakly to a preexisting protein (in this case, an aminoacyl-tRNA synthetase). Amino acid insertions in the synthetase protein that did not interfere with its synthetase activity but that enhanced its ability to interact productively with the RNA would be subject to positive selection. “Interact productively” could include increasing the rate or accuracy of splicing, allowing regulation of activity, or permitting activity under a wider variety of temperatures or other conditions. As protein took over part of the RNA-stabilizing function, portions of the RNA structure would become superfluous and would accumulate mutations and deletions (Paukstelis et al. 2008). The present-day intron RNA-CYT-18 protein RNP enzyme has been captured in a crystal structure (Fig. 2). This is presumably just today’s snapshot of an ongoing evolutionary process.

The second story concerns the evolution of the telomerase RNP. I begin by noting that the TERT protein subunit is conserved among eukaryotes and that its reverse transcriptase domain is related to those of retroviral and retrotransposon reverse transcriptases (Nakamura and Cech 1998). Thus, the ancestral telomerase was presumably a reverse transcriptase protein that copied an RNA template that was only loosely bound. Insertions and mutations in both the protein and the RNA that enhanced their ability to form a stable productive complex would be advantageous. Thus, the reverse transcriptase acquired the RNA-binding domain seen in contemporary TERTs, and the RNA acquired TERT-binding sequences. Once the enzyme was acting as a stable RNP, any further additions to the RNA that allowed recruitment of accessory proteins were subject to positive selection (Zappulla and Cech 2004). Furthermore, any changes in the RNA that supplemented the protein’s active site for DNA synthesis would also be subject to selection (Qiao and Cech 2008) and might even permit degeneration of portions of the protein. It is clear that telomerase evolution is a multi-branched pathway, because different telomerases have very different-sized RNA subunits and very different protein to RNA ratios.

**WHAT DOES THE FUTURE HOLD FOR RNPS?**

Some would argue that RNPs are on the way out as biological catalysts and that protein enzymes will assume their functions. Indeed, there are some findings that can be interpreted as supporting such a proposition, although these putative “missing links” are few.

One case is the human mitochondrial ribosome, which has a ratio of tRNA to protein much less than other ribosomes (O’Brien 2002; Sharma et al. 2003, 2009). This has led some to wonder whether the rRNAs could disappear entirely and be replaced by protein. However, before we give too much weight to this fascinating observation, it must be pointed out that mitochondria are highly specialized and are under unusual selective pressures. For example, pressure to minimize the size of the mitochondrial genome, coupled with the difficulty of importing large RNAs through the mitochondrial membrane and the relative ease of importing proteins from the cytoplasm, would provide a strong driving force for a mitochondrial ribosome with a low RNA to protein ratio. Because mitochondrial ribosomes translate only a very small number of mRNAs, a simpler RNP machine may be quite adequate. In addition, before citing the mitochondrial ribosome as an example of “disappearing RNA,” we must consider whether it could represent an ancient state rather than recent evolution. If the protein-rich mitochondrial ribosome were a fossil of a more ancient ribosome, frozen in time because of its limited functional requirements, it would represent a missing link in the acquisition of more RNA elements in the modern ribosome—exactly the opposite of the “disappearing RNA” hypothesis!

A dramatic example of disappearing RNA is the human mitochondrial RNase P. RNase P is an RNA-processing enzyme that makes a specific cut in the precursors to tRNAs, and bacterial RNase P enzymes consist of a catalytic RNA subunit supported by a small basic protein...
that aids in RNA conformational stability. RNase P enzymes of Archaea and Eukarya are also RNP s, with the RNA subunits clearly related to those of Bacteria and therefore derived from a common ancestor (Evans et al. 2006). However, the human mitochondrial RNase P is composed entirely of three protein subunits, each unrelated to any known component of RNase P in other organisms (Holzmann et al. 2008). As in the case of the protein-rich mitochondrial ribosome described above, the unusual selective pressure experienced by the human mitochondrion may be driving evolutionary events that would be unlikely to be successful at the whole-cell level.

Balanced against these rare examples that could perhaps support the view of the demise of RNPs, abundant examples support the view that in the modern post-DNA world, many new catalytic RNPs have arisen. These are of two major types. The microRNAs are involved in translational repression, acting by binding to sequences in the 3′-untranslated regions of mRNAs and in some cases by directing cleavage of the mRNA at their binding site. The active entity is an RNP consisting of the small RNA bound to a protein of the Ago (Argonaute) class. The Ago protein has in fact been shown to have an active site for mRNA cleavage (Song et al. 2004), so it is a classical RNP enzyme.

The second major role for catalytic RNPs is seen in mammals, where long noncoding RNAs bind proteins including histone-modifying enzymes that promote transcriptional silencing. The most famous example is the Xist RNA that coats only one of the two X chromosomes in female mammals, binding proteins that promote chromosome condensation and transcriptional inactivation. The end result is gene dosage compensation—the same number of active X-chromosome genes in females as in males. A rather similar situation occurs much more locally along chromosomes. For example, a long noncoding RNA is transcribed upstream of and in the same direction of Figure 2 and Jack Szostak (Massachusetts General Hospital), Paul Ktari (National Institutes of Health), and the late Jack Szostak (Massachusetts General Hospital) for comments on the manuscript.

**REFERENCES**


Pena V, Rozov A, Fabrizio P, Lührmann R, Wahl MC. 2008. Structure and function of an RNase H domain at the heart of


Evolution of Biological Catalysis: Ribozyme to RNP Enzyme

T.R. Cech

Cold Spring Harb Symp Quant Biol 2009 74: 11-16 originally published online October 22, 2009
Access the most recent version at doi:10.1101/sqb.2009.74.024

References
This article cites 41 articles, 16 of which can be accessed free at:
http://symposium.cshlp.org/content/74/11.full.html#ref-list-1

License

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.

To subscribe to Cold Spring Harbor Symposia on Quantitative Biology go to:
http://symposium.cshlp.org/subscriptions

Copyright © 2009, Cold Spring Harbor Laboratory Press