J Mol Evol (1998) 46:18–36

JOURNAL OF MOLECULAR EVOLUTION

Relics from the RNA World

Daniel C. Jeffares, Anthony M. Poole,* David Penny

Institute of Molecular Biosciences, Massey University, PO Box 11222, Palmerston North, New Zealand

Received: 14 January 1997 / Accepted: 19 May 1997

Abstract. An RNA world is widely accepted as a probable stage in the early evolution of life. Two implications are that proteins have gradually replaced RNA as the main biological catalysts and that RNA has not taken on any major de novo catalytic function after the evolution of protein synthesis, that is, there is an essentially irreversible series of steps RNA → RNP → protein. This transition, as expected from a consideration of catalytic perfection, is essentially complete for reactions when the substrates are small molecules. Based on these principles we derive criteria for identifying RNAs in modern organisms that are relics from the RNA world and then examine the function and phylogenetic distribution of RNA for such remnants of the RNA world. This allows an estimate of the minimum complexity of the last ribo-organism—the stage just preceding the advent of genetically encoded protein synthesis. Despite the constraints placed on its size by a low fidelity of replication (the Eigen limit), we conclude that the genome of this organism reached a considerable level of complexity that included several RNA-processing steps. It would include a large protoribosome with many smaller RNAs involved in its assembly, pre-tRNAs and tRNA processing, an ability for recombination of RNA, some RNA editing, an ability to copy to the end of each RNA strand, and some transport functions. It is harder to recognize specific metabolic reactions that must have existed but synthetic and bio-energetic functions would be necessary. Overall, this requires that such an organism maintained a multiple copy, double-stranded linear RNA genome capable of recombination and splicing. The genome was most likely fragmented, allowing each ''chromosome'' to be replicated with minimum error, that is, within the Eigen limit. The model as developed serves as an outgroup to root the tree of life and is an alternative to using sequence data for inferring properties of the earliest cells.

Key words: Molecular evolution — Molecular fossils — Ribozyme — RNA world — Spliceosome — Theoretical biology — Ribo-organism — Small nucleolar RNA — *Riborgis eigensis*

Introduction

Suggestions that, in early life forms, RNA was both genetic material and catalyst began with the first investigations of the ribosome (Crick 1968; Orgel 1968), long before the idea gained support from the discovery of RNA with catalytic activity (Cech 1986; Guerrier-Takada et al. 1983). This concept of an RNA world (Gilbert 1986) has been widely discussed (for example: White 1982; Reanney 1987; Watson et al. 1987; Benner et al. 1989; Gibson and Lamond 1990; Gesteland and Atkins 1993) and at present there are no serious alternatives to an RNA world being one essential intermediate stage in the origin of life. The primary evidence for an RNA world comes from the roles of RNA in modern cells; these are considered relics or molecular fossils from an earlier living system. Recent advances in the study of RNA, particularly of RNP (ribonucleic acid plus protein), have allowed much closer scrutiny of this stage in the evolution of life. Neither the origins of nor the

**Present address:* Department of Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan

Correspondence to: D. Penny; e-mail D.Penny@massey.ac.nz

energy sources available in the RNA world are discussed here; recent considerations of these problems can be found in Wächtershäuser (1990, 1992), Joyce and Orgel (1993), Böhler et al. (1995), and Lazcano and Miller (1996), along with supportive experimental data (Ertem and Ferris 1996).

Many aspects of the RNA world remain uncertain, as does the transition to a protein-encoded system. One concern is just how rapidly a protein translation system developed. In some schemes a complex DNA/RNA world is envisaged before translation (Benner et al. 1989), whilst in others some protein translation is seen as developing quite early in a relatively simple RNA world (Gibson and Lamond 1990), and more recently that protein is seen being as involved from very early stages (Gordon 1995). A related problem is the root of the tree of life. Determining the root would define many features of the Last Universal Common Ancestor (LUCA) of all living systems (Penny 1988) and allow us to distinguish between ancestral and derived biochemical features (Rivera and Lake 1992; Gupta and Singh 1994).

Our approach is to review the functions that RNA performs within modern cells and to consider the phylogenetic distribution of these RNAs. Based on the assumption that *ubiquitous* classes of RNAs are molecular fossils of the RNA world, we describe a model (*Riborgis eigensis*) for the structure of ribo-organisms at the point where genetically encoded protein translation evolved: the breakthrough organism. This model will necessarily be a conservative estimate of the final complexity of the RNA world because many original ribozymes will have been replaced by proteins during the long time period since the evolution of translation.

Several discussions (Diener 1993; Maizels and Weiner 1993, 1994; Gibson and Lamond 1990; Reanney 1987) consider the possible genome structure of ribo-organisms pre-dating the three primary lineages (eukaryotes, archaea, and eubacteria) recognized today. ''Genome organization'' in this discussion includes only information such as the size of genome, whether it is linear or circular, fragmented or continuous, the copy number, and the presence of single or multiple centers of origin of replication. It does not include cytological information such as cellular or acellular, membranes present or absent, or the degree, if any, of cellular compartmentation (such as a nuclear structure).

We consider genome architecture fundamental, and because archaea and eubacteria are fundamentally similar in genome organization (Baumann et al. 1995), we distinguish between them only as necessary, though this does not prejudge the question of the monophyly of the prokaryote genome organization (Forterre 1996). An important question is whether a circular or linear genome architecture is more plausible for the ancestral condition. In agreement with others we show, based on genome dynamics and the molecular fossil record, that a linear, recombination-competent, fragmented genome is the most probable ancestral state (Fig. 4). Finally, we set tentative upper and lower size limits for the genome that are based on the inferred complexity of metabolism in this model ribo-organism.

To be useful, a scientific model should lead to predictions and in the present case we test our model to see whether it could lead to later stages in the evolution of life. The origin of life involves a continuous series of ancestors, all of which need to be fully viable. The stage after the last ribo-organism is the evolution of protein biosynthesis. By examining those RNAs that have survived until the present day we demonstrate (Poole et al. 1997) that an encoded translation system could develop by numerous small steps from our model ribo-organism. Finally we use this model to root the tree of life by examining RNA metabolism as an alternative to using sequence data. A conclusion is that the prokaryote genome is derived in structure, having undergone *r* selection and/or a thermophilic stage, becoming smaller, more compact, and efficiently organized. An interesting corollary of the analysis is that the genome structure of eukaryotes most closely approximates the architecture expected for the genome of the last ribo-organism (Jeffares et al. 1995; Poole et al. 1997).

Basic Assumptions

We developed our model for the final complexity of the RNA world based upon several assumptions and evolutionary principles. The key ones are given below and additional information is given later.

1. *An RNA world existed and ribo-organisms were the immediate predecessors of translation-competent organisms.* The unique ability of RNA to act as information carrier and catalyst provides strong support for the existence of the RNA world. Many functions of the ribosome are RNA-mediated (Noller 1993a) and the peptidyl transferase function can be performed by protein-depleted ribosomes (Noller et al. 1992), indicating that protein synthesis arose in ribo-organisms.
2. *After the evolution of translation, proteins have augmented, then largely replaced RNA as catalysts.* Catalytic RNA took on no major de novo functions after the evolution of genetically encoded protein translation. We envisage a double transition, first to an RNP (ribonucleic acid plus protein) world, and later to a DNA-protein world with protein carrying out nearly all catalytic functions. New RNAs or RNPs could possibly arise by gene duplications and divergence to a modified function, but this seems to be very rare.
3. *Frozen accidents are a consequence of evolution as tinkering.* The concept of a ''frozen accident'' (Crick 1968; see also Eck and Dayhoff 1966), describes an

evolutionary event that is retained not because it is the optimal state but because many other systems rely on it. The concept is important in accounting for RNAs being retained over the 3 billion or more years since the evolution of translation. Maizels and Weiner (1994) call this concept ''coevolutionary dependence.'' We find that many of the RNAs in modern organisms are central to metabolism, and under assumption 2 they are not expected to arise in a modern, protein-synthesizing organism. Moreover, ribozymes not central to metabolism are more likely to be replaced by proteins.

4. *Relic RNA can be defined according to simple criteria.* Not all RNAs need be molecular fossils dating to the RNA world. Here we use three criteria to evaluate the antiquity of RNAs. Many of the RNAs satisfy all three criteria, which are detailed below:

   A. Catalytic. Catalytic RNAs are more likely to be relics than are information-carrying RNAs. The advantage of proteins over RNA lies in their superiority as catalysts. Hence we may assume such RNAs are not recent additions to metabolism. Where RNA acts solely as an information carrier, and could have arisen de novo or by duplication and divergence from earlier RNAs, the case for antiquity is not strong. RNA could even have advantages that would be selected over protein—for example, antisense RNA could easily arise from double-stranded DNA and could have intrinsic advantages as a gene regulation mechanism.

   B. Ubiquitous. RNAs that occur in all living organisms are more likely to be ancient molecular fossils because it is unlikely that novel RNAs will arise independently in several lineages. However, the decision as to whether an RNA is ubiquitous will change with different rootings of the tree of life.

   C. Central to metabolism. Possibly the most important consideration is the position of an RNA molecule in metabolism. Under the frozen accident concept (assumption 3) any RNA molecule central to the workings of the cell is very difficult to replace and thus more likely to date from the RNA world.

5. *Continuity of function—evolution by tinkering.* Under a Darwinian mechanism any complex structure cannot arise by chance de novo; there must be many functional intermediates. Evolution ''tinkers'' (Jacob 1977) with what is already present. It is possible some RNAs, in particular those which are not central to metabolism, have been altered so that their functions now differ from those in the RNA world.

6. *Eigen limit.* The quasi-species concept of Eigen and Schuster (1977; Eigen and Winkler-Oswatitsch 1981; Eigen 1993; see Maynard Smith and Szathmáry 1995, p44) demonstrated that the accuracy of replication placed limits on the size of genome that can be maintained by selection. The higher the error rate during replication, the smaller the maximum permissible genome size. We assume that replication accuracy was a strong limiting feature in the RNA world.

Table 1. Turnover numbers for ribozymes and proteins[a]

Catalyst	k_{cat} (min^{-1})	k_{cat}/K_m (M^{-1} min^{-1})
Tetrahymena L-21(SacI)	0.1	9.0×10^7
Polynucleotide kinase[c]	0.3	6.0×10^3
19-base virusoid	0.5	8.3×10^5
L-19 intron	1.7	4.3×10^4
RNase P RNA	1	2.0×10^6
RNase P RNA and protein	2	4.0×10^6
RNase T1[b]	5,700	1.1×10^8
Staphylococcal nuclease[b]	5,700	6.0×10^8
T4 polynucleotide kinase[b]	25,000	6.0×10^8
Triose-P isomerase[b]	258,000	1.4×10^{10}
Cyclophilin[b]	780,000	9.0×10^8
Carbonic anhydrase[b]	600,000,000	7.2×10^9

[a] Turnover number is k_{cat}. Values from Herschlag and Cech (1990); Lewin (1994; p953); Guerrier-Takada et al. (1983); Lorsch and Szostak (1994); Radzicka and Wolfenden (1995); Lillehaug and Kleppe (1975)
[b] Protein catalysts
[c] Artificial ribozyme evolved in vitro

The Replacement of RNA by Protein

A consequence of the RNA world model is that, over time, proteins have gradually replaced RNAs in catalysis by virtue of their superior catalytic properties (Orgel 1968; Crick 1968; Watson et al. 1987); protein catalysts generally have much faster turnover numbers and shorter reaction times (Table 1). The expectation is a two-stage transition of RNA ribozymes to proteins via ribonucleoprotein ribozymes (RNP), that is:

$$\text{RNA} \rightarrow \text{RNP} \rightarrow \text{protein}$$

and consequently no novel catalytic RNAs are formed de novo after the advent of efficient genetically encoded protein synthesis. We consider each step in turn: RNA ribozymes to RNP ribozymes and RNPs to protein enzymes, though for this latter phase we distinguish small molecule substrates from RNA substrates.

The selection pressures that led to the almost exclusive usage of protein enzymes in modern metabolism are only partially understood, especially as many cofactors appear to date from the RNA world (White 1982), and at least uracil could have had amino acids covalently linked (Robertson and Miller 1995b). The number of functional groups (20 amino acids) used by proteins, and the ability to maintain precise tertiary structure, is an advantage which RNA catalysts lack. Proteins are able to perform many of the functions that RNAs still catalyze in nature—RNase E (a protein) can cleave RNA in much the same way as the ribonucleoprotein RNase P (Apirion and Miczak 1993) or hammerhead ribozymes, and catalytic antibodies can form peptide bonds (Jacobsen and Schultz

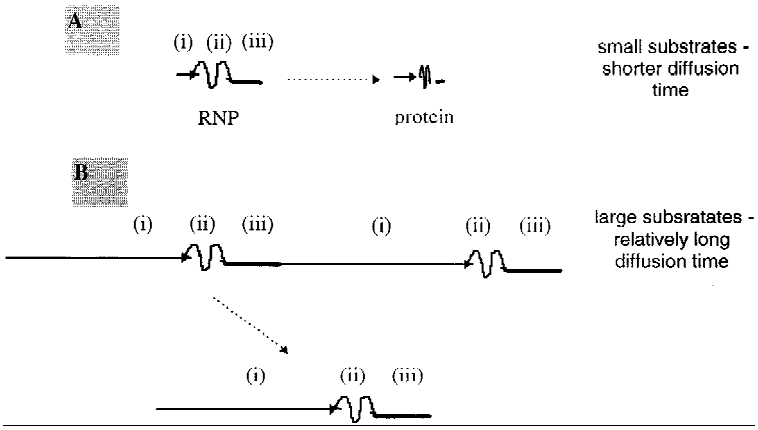

Fig. 1. Comparison of reaction rates for small and large substrates. Each enzymatic reaction is idealized as a ''diffusion'' time when the substrate(s) and catalyst meet (*i*), a ''reaction'' time (*ii*), and a ''release/readjustment'' time before the catalyst is free to begin the next reaction (*iii*). Given Graham's Law of Diffusion (rate inversely proportion to $\sqrt{MW}$), the diffusion time for a large substrate will be much longer than the reaction time for the catalyst. Shortening the reaction time would not necessarily give much advantage with a large (macromolecular) substrate. An RNP may suffice in cases where diffusion limits the multiple turnover of products. In such a situation the only solution would be to remove the need for the entire process. The differences in the selection pressures of prokaryotes (relatively *r* selected) and eukaryotes (relatively K selected) could account for the large variations in the numbers of RNPs used.

1994), a reaction that appears to be RNA catalyzed in the ribosome (Noller et al. 1992; Noller 1993b). Nor is there any need for RNA to recognize any specific RNA sequence, because a variety of proteins bind RNA in a sequence-specific manner (Draper 1995). Thus recognition of specific RNA sequences by RNA in the spliceosome (Madhani and Guthrie 1994b), and in the tRNA-ribosome interaction (von Ahsen and Noller 1995; Samaha et al. 1995) could, in theory, be replaced by proteins. This observation, given our assumption that RNA does not take on any new function after the evolution of protein synthesis, implies such RNA-RNA recognition systems would date from the RNA world.

RNA → RNP

In vitro evolution experiments have shown that RNA by itself can form remarkably efficient catalysts. Herschlag and Cech (1990) describe a ribozyme from *T. thermophila* derived by artificial selection with a $k_{cat}/K_m = 9 \times 10^7$, a value of an order comparable to that for protein catalysts (see Radzicka and Wolfenden 1995). Ribozymes can also be highly selective about their substrates; hammerhead ribozymes and self-splicing introns display high specificity from using internal guide sequences (Cech and Uhlenbeck 1994; Pyle et al. 1992). In extant life ribozymes catalyze transesterification reactions but in vitro evolution experiments show that RNA is able to catalyze many other reactions (e.g., Doudna et al. 1991; Bartel and Szostak 1993; Lorsch and Szostak 1994; Wilson and Szostak 1995; Illangasekare et al. 1995). However very few catalysts are formed entirely of RNA in extant organisms—self-splicing introns, viral hammerhead ribozymes, and related self-splicing RNAs such as the Varkud retroplasmid (Maizels and Weiner 1993) are exceptions.

Despite the catalytic prowess of ribozymes, evidence to support the first step from RNA to RNP abounds. The RNase P holoenzyme (RNA plus protein) cleaves tRNA-like substrates more efficiently than the RNA subunit alone and requires lower Mg^{2+} concentrations (Liu and Altman 1994). Although it could be argued that the RNA has atrophied in the presence of its protein moiety, this in itself is indicative of the direction of evolution. A limitation on catalysis by ribozymes appears to be in RNA forming and maintaining an optimal tertiary structure, as evidenced by requirement for high concentrations of Mg^{2+}—the RNase P RNA, for example (Guerrier-Takada et al. 1983). Chaperone-like activities of nonspecific proteins may help maintain the active site and help neutralize negative charges on the RNA (see later). Nonspecific RNA-binding proteins can enhance the activity of ribozymes that are not associated with such proteins in nature (Herschlag et al. 1994; Coetzee et al. 1994). It thus appears that ribozyme stabilization and enhancement by peptides does not require highly specific amino acid interactions, making the transition to RNP easier and, given the increased activity referred to above, advantageous.

RNP → Protein

This transition is considered separately for small and large substrates because differences are expected depending on whether or not diffusion of the substrate onto the catalyst is rate-limiting (Fig. 1). For small substrates there are now apparently no RNP ribozymes, though in a few cases an amino acid bound to a tRNA is involved in intermediary metabolism. These include: glutamate to glutamate-1-semialdehyde conversion in the early stages of heme synthesis in chlorophyll (Schön et al. 1986), the use of $tRNA^{Gly}$ in *Staphylococcus* peptidoglycan synthesis (Roberts et al. 1968); the ''mischarging'' of $tRNA^{Gln}$ with glutamate and the subsequent conversion of the bound glutamate to glutamine (Rogers and Söll 1995); and the similar mischarging of $tRNA^{Asn}$ with aspartate

(Curnow et al. 1996). Because these reactions are neither ubiquitous nor use a catalytic RNA moiety they do not meet our criteria for being relics of the RNA world. We suggest (Poole and Penny in preparation) that amino acid mischarging of tRNA may be a biochemical adaptation to extreme conditions when, for example at high temperatures, free glutamine is unstable and its low concentration would limit protein biosynthesis. In the present context, the examples above serve to emphasize that catalysis involving small molecules is essentially carried out by proteins, albeit often with a ribonucleotide cofactor (NAD, NADP, FAD, etc.).

When the substrate is a large macromolecule, the replacement of RNP by protein is not as complete as for small substrates; this may be explained by the relative rates of catalysis and substrate diffusion. An enzyme is deemed to have reached catalytic perfection when diffusion of substrate onto the active site is slower than the rate of chemical catalysis, and the diffusion step is thereby rate-limiting (Albery and Knowles 1976). Under such circumstances, refining the chemical step will not enhance the overall rate of catalysis; thus the enzyme can be considered to have reached evolutionary perfection. In the case of small molecules, diffusion rate is relatively fast, and in many instances the rate of the chemical reaction, not diffusion, is rate-limiting. In such cases refinement of the active site will improve the overall catalytic rate because the efficiency of the enzyme is dependent on the chemistry of the reaction. As the chemistry afforded by protein catalytic sites is superior to that of RNA catalytic sites (Table 1), we expect that there is a tendency for protein enzymes to replace RNP ribozymes when the chemical step, and not diffusion, is rate-limiting (Fig. 1A).

However, in the case of large substrates such as RNA the rate of diffusion is the slowest step and so refinement of the chemical step will not speed the reaction overall (Fig. 1B). Thus catalytic perfection and the efficiency of the chemical step cannot be equated—catalytic perfection is reached when the overall rate of catalysis is limited only by the diffusion of the substrate onto the active site. This explains the observation that extant RNP ribozymes appear to carry out enzymatic catalysis exclusively on large macromolecule RNAs. In this case catalytic perfection can be reached without a protein-catalyzed step because it is the diffusion of substrate onto the catalyst that is limiting the overall rate of the reaction (diffusion being inversely proportional to $MW^{0.5}$ —Graham's Law of Diffusion). Consequently, there is little selection to replace the RNA enzyme, irrespective of the fact that protein chemistry is superior.

Phylogenetic Distribution of RNAs

Given the strong trend RNA → RNP → protein, we can examine the phylogenetic distribution of RNA in modern organisms for relics and thus develop a model for complexity in the RNA world (Benner et al. 1989; White 1976, 1982; Yarus 1993; Gibson and Lamond 1990; Diener 1993). Candidates are RNAs which fit at least one of the criteria: catalytic, ubiquitous (or at least conserved within the eukaryotic lineage, see later), or central to some aspect of metabolism. These are summarized in Figs. 2 and 3 and show the distribution and sizes of the RNA moieties, thus allowing estimates of the minimum genome size of the final ribo-organism. We will review each major class in turn.

Ribosomes

Although peptide synthesis was likely before the evolution of translation—even in prebiotic environments (Lahav et al. 1978; White and Erickson 1980)—protein structure must be inherited before it can evolve. Many of the ribosome functions required for protein synthesis are, at least in part, RNA-mediated, including peptidyl transferase (Noller et al. 1992) and decoding functions (Purohit and Stern 1994; von Ahsen and Noller 1995), and the tRNA acceptor site interaction with 23S rRNA (Moazed and Noller 1991). In addition, prokaryotic ribosomes depleted of protein will catalyze the fragment reaction, and thus even modern ribosomes appear to be protein-stabilized ribozymes (Noller et al. 1992; Noller 1993b). This result does not, however, rule out the possibility that proteins play an essential role in the catalytic function of the modern ribosome, and to this effect it is not surprising to find that the ribosomal protein L2 is essential for peptidyl transferase activity (Cooperman et al. 1995) in some organisms. This finding provides evidence for the reality of the RNA-to-protein catalytic transition over time in that the ribosome has taken on the superior catalytic efficiency of protein chemistry.

Decoding, the interaction of tRNA anticodons with the ribosome, can be mimicked by a small RNA analog of the rRNA region thought to be involved in decoding in intact ribosomes (Purohit and Stern 1994). Smaller RNAs could thus have acted in *trans* to form a functioning ribosome, in much the same way as the larger 5S, 16S, and 23S function today. Additional evidence for a composite protoribosome may be the finding that the α-sarcin loop appears to be a modular RNA (Szewczak and Moore 1995). Still further evidence which makes the concept of an RNA-only ribosome feasible is the demonstration that an in vitro-evolved RNA is capable of tRNA aminoacylation (Illangasekare et al. 1995). These points, taken together, demonstrate that although proteins are now essential for high turnover in amino acid synthesis, an all-RNA ribosome is now, more than ever before, a plausible part of the RNA world model.

How is it that the protoribosome evolved to such complexity with RNA, and so many of its functions are still carried out by RNA? A possible answer to both questions

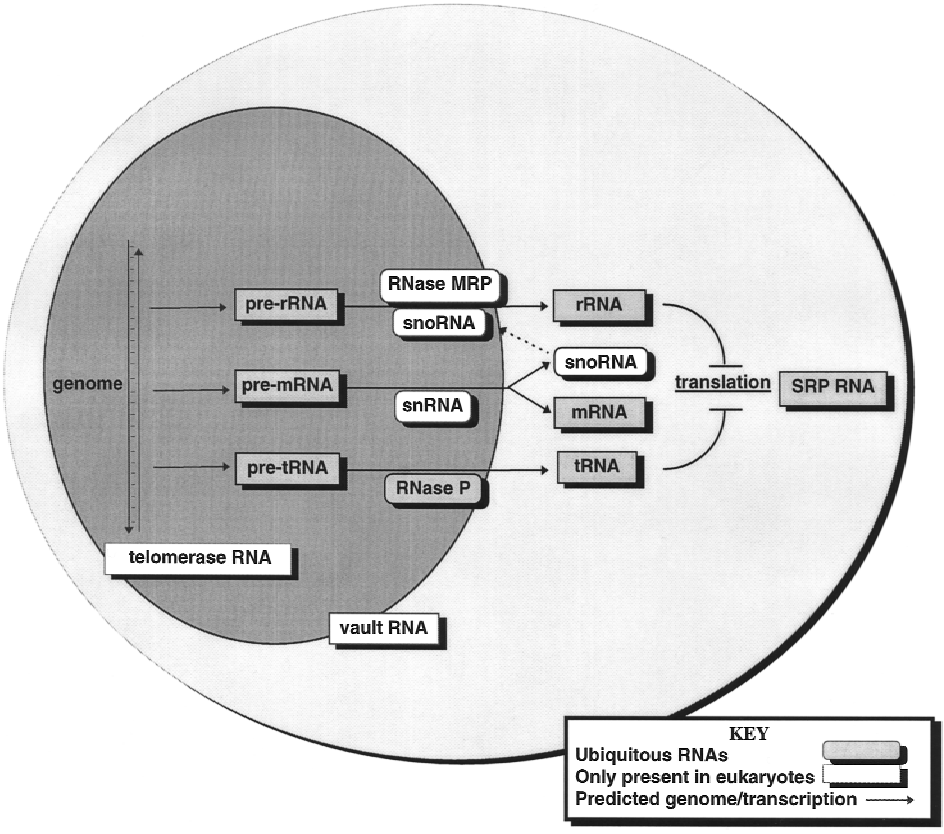

Fig. 2. RNAs involved in the processing of RNA in modern cells. Processing of pre-rRNA, pre-mRNA, and pre-tRNAs shows a similar order of RNP-mediated steps in transcript maturation, though those shown in *white* have been lost (on our interpretation) in prokaryotes—although at least one archaeon (*Sulfolobus acidocaldarius*) contains a snoRNA (Potter et al. 1995). It is also interesting that all the universally conserved RNAs (ubiquitous) contribute in some way toward the protein translation apparatus. Note that the picture includes a nucleus. This we find useful to include only because it demarcates the boundary between the processes of transcription and translation.

may come from studying problems in a cyclic system with low fidelity. Hasegawa et al. (1984) showed that the components of the translation apparatus are subject to error amplification if they synthesize their own constituent proteins. The basic concept of amplification of errors in a cyclic system was introduced by Orgel (1963) in the context of aging. Any errors present in such a system that decrease fidelity are amplified and the whole system degrades in an ''error cascade'' (Eigen and Schuster 1979) or ''mutational meltdown.'' An early, low-fidelity protein catalyst would initially increase error amplification; the first proteins are likely to have been selected for their ability to increase the accuracy of the existing protoribosome. This argument is independent of the origin of the protoribosome because whether it was involved in replication or ribogene transcription, it was susceptible to an error cascade (Hasegawa et al. 1984). We prefer the model in which the protoribosome initially had a different function (possibly originating in replication, Gordon 1995) and it was only later recruited into the role of protein synthesis. A detailed model is possible (Poole et al. 1997). The estimate we use for the size of the original ribosome is approximately 7,500 nucleotides (Table 2), which includes both major subunits, 5.8S rRNA, and intervening sequences.

Snorposome—Processing and Assembly of Ribosomes

The eukaryotic rRNA maturation apparatus includes many ribonucleoproteins from the ever-broadening family of small nucleolar RNA (snoRNA) (Table 2; Morrissey and Tollervey 1995; Maxwell and Fournier 1995). These RNAs differ from small nuclear RNA (snRNA) and appear to constitute a larger functional unit, known as the snorposome or processosome, which is involved in ribosome maturation. Many snoRNAs are encoded within introns of proteins such as HSC70 (Liu and Maxwell 1990; Leverette et al. 1992), RCCI (Kiss and Filipowicz 1993), and a number of ribosomal proteins (reviewed in Maxwell and Fournier 1995; Table 2). The most striking example of this is eight different small snoRNAs encoded in eight different introns of a single gene (Tycowski et al. 1996). The total number of snoRNAs is not yet known but at least 30 appear to occur in mammals (Kiss-László et al. 1996; Nicoloso et al. 1996). These snoRNAs may turn out to be some of the most striking remnants from the RNA world.

The first indication of their function was revealed as a result of finding that parts of their sequences were complementary to parts of the RNA of ribosomes, including the 5′ETS, 18S rRNA, ITS, 5.8S rRNA, and 28S

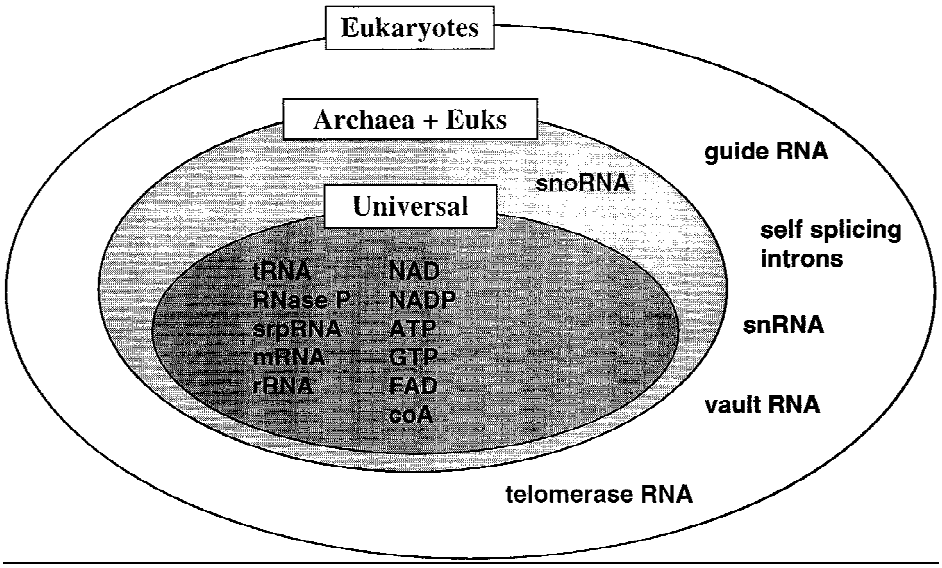

Fig. 3. Phylogenetic distribution of molecular fossil RNAs that may date back to the RNA world. In this Venn diagram, RNA includes pyridine cofactors possessed by all living groups (eubacteria, archaea, and eukaryotes) as these are central to metabolism. Eukaryotes contain all the RNAs shown and hence represent the superset. Prokaryotes, in particular eubacteria, lack several RNPs that our model predicts originated before protein translation. We suggest that prokaryotes have lost more RNAs than eukaryotes and that this loss may be due to the relative influence of *r* selection and/or thermophily.

rRNA (Maxwell and Fournier 1995), in some cases to conserved single-stranded regions. Crosslinking studies indicated that both U3 and snR30 (Table 2) are in contact with the 35S pre-rRNA in vivo (Morrissey and Tollervey 1993) and the base pairing between U3 and the pre-rRNA can be widely conserved (Beltrame and Tollervey 1992). For example, Peculis and Steitz (1994) found that many bases in the 5′ region of *Xenopus* U8 are essential for rRNA processing and that this region is highly conserved between *Xenopus* and mammals (90% identity). The intron-exon structure of the region encoding snoRNA (Sollner-Webb 1993) may also be conserved. There are also potential base-pairing interactions with core regions of rRNAs (Bachellerie et al. 1995) and some intron-encoded snoRNAs appear to require spliceosome processing before final maturation (Kiss and Filipowicz 1995). All these support the antiquity of the ribosome, spliceosome, and snorposome system. Some snoRNAs (U32, U36) have conserved regions complementary to both the 18S and 28S rRNAs (Nicoloso et al. 1996) and so a single snoRNA molecule may form base pairs with both rRNAs. As well as a potential role in assembly of the rRNA complex, this may be indicative of an ancient RNA scaffold that maintained rRNA conformation both intra- and intermolecularly where there is now a protein scaffold. It is significant that this offers a solution to the problem of assembling large RNA complexes before the advent of proteins.

An important function of snoRNAs is to identify nucleotides where the 2-hydroxy position of ribose becomes methylated during the maturation of the ribosome (Kiss-László et al. 1996; Nicoloso et al. 1996). Methylation occurs five nucleotides from the paired region between the snoRNA and the rRNA. Direct experimental support for this role in determining sites of methylation is now available from genetic studies (Cavaille et al. 1996). The role of ribose methylation is not yet clear but is suggested to increase hydrophobicity, allowing closer packing of the eukaryotic ribosome. The absence of snoRNAs from prokaryotes (apart from some archaea, see later) is an enigma (see Poole et al. 1997, for a possible solution) and the virtual absence of methylated rRNA in prokaryotes (Nicoloso et al. 1996, and references therein) indicates that extensive methylation is not essential for ribosome action, at least once protein synthesis is established. It can only be assumed at present that in prokaryotes, proteins fulfill the role that methyl groups perform in eukaryotes. Assuming RNA → RNP → protein, this use of snoRNA would be the ancestral condition (see Table 3 in Poole et al. 1997).

A functional homologue of the eukaryotic U3 snoRNA has been found in the archaeon *Sulfolobus acidocaldarius* (Potter et al. 1995). In contrast, the processing of pre-rRNA in eubacteria appears to involve only proteins. Thus it is important to examine rRNA processing in other archaea. Nevertheless, finding U3 snoRNA in eukaryotes and an archaeon supports a gradual replacement of snoRNPs by protein enzymes in prokaryotes (Jeffares et al. 1995), and we discuss the selective advantages of this loss of RNPs later (Poole et al. 1997). The total length of the 26 snoRNAs of yeast is 5,426 nucleotides (Maxwell and Fournier 1995) and we will use this as our estimate of the size of the ancestral snorposome and related functions.

Table 2. RNA function in modern cells

Function	RNA molecule	Size (bp)	Distribution	In RNA world?	Reference/Genbank accession no.
Translation: structural and catalysis	16S/18S rRNA (small subunit)—eukaryote 18S is equivalent to prokaryote 16S	*E. coli* 1542 *Xenopus* 1826 ITS1 557	Ubiquitous	Yes	J01695; X59734
	23S/5.8 & 28S rRNA (large subunit)—eukaryote 28S + 5.8S is equivalent to prokaryote 23S	*E. coli* 2904 *Xenopus* 5.8S 162 ITS2 262 28S 4115 ETS 712	Prokaryotes Eukaryotes Eukaryotes	Yes	J01695 X59734 X59734
	5S rRNA	*E. coli* 120 *Xenopus* 193	Ubiquitous	Yes	J01695; M10850
	rRNA gene cluster	*Xenopus* rRNA genes 7634 *E. coli* rrnB operon 7508			X59734 J01695
Translation, information carrier	mRNA	Various	Ubiquitous	At beginning of RNP world	
Translation: adaptor	tRNA	*E. coli* mature tRNA 76 (supB-E tRNA cluster with 7 tRNAs 1100)	Ubiquitous	Yes	J01713
Amino acid metabolism	tRNA		Ubiquitous	Yes	
Replication primer	tRNA		Retrotransposons, Retroviruses, Plant pararetroviruses	Possibly	
tRNA maturation	RNase P	*E. coli* 377	Ubiquitous	Yes	M17569
rRNA processing	snoRNAs 26 total *S. cerevisiae.*	5426 total	Eukaryotes	Yes	Maxwell and Fournier (1995)
	U3 snoRNA		Sulfolobus acidocaldarius (an Archaeon)		Potter et al. (1995)
rRNA processing	RNase MRP	*Xenopus* 277	Eukaryotes	Related to RNase P	D30745
mRNA splicing in eukaryotes	Spliceosomal RNAs	*Xenopus* U1 122 U2 189 U5 116 Yeast U4 129 U6 112 Human U4atac 131 U6atac 125	Eukaryotes	Yes	 K02697 K02457 X06020 X15491 X12565; X54302
mRNA splicing	Group I self-splicing introns	*Tetrahymena*		Possibly	
mRNA splicing	Group II self-splicing introns			Possibly	
Mistranslated protein degradation	10Sa RNA	*E. coli* 363	Eubacteria	Possibly, though RNA world function unknown	D12501; Tu et al. (1995); Keiler et al. (1996)
Modulation of DNA-binding protein activity	10Sa RNA			Nonspecific chaperone-binding modulator in RNP world?	D12501; Retallack and Friedman (1995)

Table 2. Continued.

Function	RNA molecule	Size (bp)	Distribution	In RNA world?	Reference/Genbank accession no.
Protein translocation	srpRNA	*Xenopus* 297 *Solfolobus solfataricus* 309	Eukaryotes, Archaea (7S RNA)	Yes (probably a GTPase)	K01754; Kaine (1990)
		E. coli 114	Eubacteria (4.5S RNA)		X01074
Maintenance of translation apparatus at elevated temperatures	G8 RNA	*Tetrahymena thermophila* 306	Eukaryotes	Probably—weak homology to 7S and 4.5S RNAs	X57037; Fung et al. (1995)
Not known	Vault RNAs	*Rana catesbeiana* 89–94, *Rattus norvegicus* 141	Eukaryotes	Possibly (information still lacking)	Z11765; Z11771
Phosphorylation	RNA nuclear extract essential for protein phosphorylation.		Eukaryotes	Possibly—speculatively an RNA-containing enzyme	Fung et al. (1997)
Gene regulation	Antisense RNA		Ubiquitous	Possibly—in gene regulation.	See text
Editing via deamination	Internal guide seqs (e.g., in tRNA)		Eukaryote mitochondria (e.g., in *Acanthamoeba castellanii*)	Yes	Lonergan and Gray (1993)
Eukaryotic chromosome end synthesis	Telomerase RNA	*Tetrahymena thermophila* 159	Eukaryotes	Possibly	Greider and Blackburn (1989)
Various	Hammer heads, viroids		Plant satellite RNA	In RNA world (see text)	
Total sizes in base pairs	**Spliceosome, snoRNAs, rRNAs, 10 tRNAs**	**7,600 (minimum)**			
	As above, with modern rRNA, tRNA gene clusters, telomerase RNA, srpRNA, vRNA	**12,000**			
	As above, including functions from table 3	**15,500**			

tRNAs and tRNA Processing—RNase P

Transfer RNAs are clear relics of the RNA world; they are ubiquitous, very highly conserved, and central to the metabolism of all life. In fact, tRNAs may be one of the oldest gene types (Eigen and Winkler-Oswatitsch 1981; Eigen et al. 1989). The direct interaction of tRNAs with rRNAs, such as the base pairing between the CCA end of tRNA and the 23S rRNA (Samaha et al. 1995), supports tRNAs pre-dating translation (Noller 1993a). The use of tRNAs as adapters for amino acid synthesis and modification (see RNP → Protein section above) and primers for the replication of viral (RNA) genomes has led to suggestions that tRNAs were used for a variety of roles in the RNA world. The genomic tag hypothesis (Maizels and Weiner 1987, 1993, 1994) presents arguments for the involvement of tRNA, or its precursors, in replication before its incorporation into the protein synthesis machinery.

During tRNA synthesis, maturation is carried out by a ribonucleoprotein, RNase P. This enzyme is interesting not only in that the RNA can be catalytic per se (Guerrier-Takada et al. 1983) but also in that the processing of pre-tRNA transcripts by RNase P appears conserved throughout the three domains (Brown and Pace 1992; LaGrandeur et al. 1993). All evidence suggests that the RNA subunit (M1 RNA) of RNase P is a molecular fossil dating from the RNA world: the M1 RNA is catalytic (Guerrier-Takada et al. 1983); the structure and function of RNase P is conserved and it occupies a central position in metabolism. Many eukaryotes possess a second RNase P-like molecule, namely RNase MRP, but it may not be an RNA world relic. It may result from a duplication/divergence event within eukaryotes (Morrissey and Tollervey 1995), or be a second copy from the endosymbiotic eubacterium that became the mitochondrion. If *Giardia* and *Microsporidia* have always lacked mitochondria (Kamaishi et al. 1996) they would, on the sec-

ond hypothesis, also lack MRP. Although the extant function of tRNA is clearly an RNA world relic, the number of tRNAs is difficult to estimate because they may have fulfilled many roles in the RNA world. An estimate of 1,000 base pairs (including RNase P) is conservative.

The Spliceosome

There is increasing evidence for catalytic RNAs in another RNP complex, the spliceosome, of which there are two versions (Tarn and Steitz 1996) depending on the consensus at the ends of the intron. Most work has been done on the common form with a GU-AG consensus, and consequently we concentrate on that. Recent UV crosslinking and genetic suppression studies have indicated that the active site of the spliceosome includes RNA structures that are similar to both group II self-splicing introns (Wise 1993; Sharp 1994) and to hammerhead ribozymes (Yang et al. 1994). RNA-RNA interactions play an important role in both the assembly of the spliceosome (reviewed in Madhani and Guthrie 1994b) and in the formation of its active site(s) (Lesser and Guthrie 1993; Sun and Manly 1995; Madhani and Guthrie 1994a), which becomes complete only with the formation of U2-U6 helices (Lesser and Guthrie 1993). Based on our criteria for relics of the RNA world, such a complex association of RNAs is consistent with the spliceosome originating in the RNA world. Because of the similarities to group II self-splicing, both in the reaction catalyzed and in structure, there has been debate as to whether these structures are related by descent or are the products of convergent evolution (Cech 1986; Madhani and Guthrie 1994b; Weiner 1993). For our purposes, the question of whether these splicing activities are evolutionarily related or are the products of ''chemical determinism'' (Weiner 1993) is less important than the antiquity of the process.

A pretranslation origin for the spliceosome seems paradoxical at first because there appears to be no need to splice mRNAs before the advent of translation, and one of our basic assumptions is that large complex structures cannot arise in the absence of selection. There are several possibilities for the protospliceosome. Given the selective advantage conferred by recombination as a means of overcoming the Eigen limit on genome size (Reanney 1979) it would be highly advantageous to have RNA splicing activity in the RNA world, perhaps as a recombinase. In support of this possibility, *trans*-splicing of ssRNA can be achieved by the spliceosome with downstream RNA sequence elements (Chiara and Reed 1995; Ares and Weiser 1995) and in addition has been performed by a single *cis*-acting RNA catalyst evolved in vitro (Sullenger and Cech 1994). Given that some snoRNAs, which are very likely to have originated in processing of the ribosome in the RNA world (Jeffares et al. 1995), are processed exclusively from spliced introns (Kiss and Filipowicz 1995), it appears that snoRNAs and the spliceosome are interdependent. Another role for RNA splicing activity could have been in the processing of RNA transcripts destined to become protoribozymes; a similar role exists in contemporary eukaryotic metabolism. An estimate of the size of an RNA world spliceosome, based just on snRNA sizes in Table 2, is about 600 nucleotides.

Signal Recognition Particle and srpRNAs

The RNA component of the signal recognition particles of eubacteria (4.5S RNA) and of eukaryotes (7S RNA) have similar structure and function (Miller et al. 1994), and some archaea also contain a similar 7S RNA (Kaine 1990). Under the assumption of RNA taking on no new function, srpRNA is expected to have existed in the RNA world, though its precise role either in the RNA world or now in protein translocation is unclear. In eukaryotes it is apparently required in the signal recognition particle (SRP) in order to stimulate hydrolysis, though not binding, of GTP (Miller et al. 1993). The homologous *E. coli* system may follow the same set of bounds; binding of the protein analogous to SRP54 to the SRP receptor requires the RNA moiety (Miller et al. 1994). It is quite feasible that the srpRNA represents an ancient ribozyme capable of GTP hydrolysis. Modern srpRNA may be an RNA in the process of being made redundant by its newer protein counterparts; only its capacity to facilitate GTP hydrolysis remains. Our size estimate is 300 nucleotides (Table 2).

RNA Editing

Several phenomena are included under the title *RNA editing;* these include nucleotide modifications and post-transcriptional insertions and/or deletions, sometimes involving guide RNA. Editing involving purine/purine, pyrimidine/pyrimidine, and pyrimidine-to-purine conversions is observed in a variety of eukaryotes (Lonergan and Gray 1993; Bass 1993; Walbot 1991). Being widespread, these modifications in tRNA and rRNA (Bass 1993) as well as in spliceosomes (Ares and Weiser 1995) are candidates for an RNA world. However, because any given form of editing appears to occur sporadically it is unclear if any are truly ubiquitous. Several forms of RNA editing occur in mitochondria, and a reasonable explanation is that organelles are susceptible to the fixation of slightly deleterious mutations because of the lack of recombination (Andersson and Hughes 1996). Mutations cannot be removed by recombination with good copies, an example of Muller's ratchet in operation (see later). It is also uncertain whether RNA editing is a simple modification of a genuinely ubiquitous step in RNA processing. Recent progress in understanding the

mechanism of insertion/deletion editing in trypanosome mitochondria will help in addressing this question (Seiwert et al. 1996; Cruz-Reyes and Sollner-Webb 1996; Kable et al. 1996).

In *Acanthamoeba* mitochondria (Lonergan and Gray 1993) it appears that four of five tRNAs in each of two clusters undergo editing by base conversion involving a guide sequence internal to the tRNA. The proposed mechanism is base conversion opposite a pyrimidine in the 3′ half of the acceptor stem, occurring in the first three bases of the 5′ part of the acceptor stem, a region involved in specific interactions with aminoacyl tRNA synthetases (Lonergan and Gray 1993; Bass 1993). In the absence of editing, appropriate secondary structure is not formed, preventing the maturation of the tRNA by RNase P. In the RNA world, this kind of editing may have, and may still, offer a tRNA expression control mechanism, providing a way of controlling tRNA molecule populations. It might also represent an early amino acid synthesis pathway or perhaps was in wider use as a means by which to control the replication system (see Table 2). A further possibility is that base conversion may represent the remains of an early nucleotide synthesis pathway. However, without adequate information it is still unclear whether RNA editing occurred in the RNA world. Hence we have not included a size estimate for RNA editing in Table 2.

Telomerase

The problem of replicating a linear nucleic acid molecule right to the end occurs in both RNA viruses and eukaryotic chromosomes (Maizels and Weiner 1994). It is overcome in eukaryotic cells by the RNP telomerase (Blackburn 1992; Cohn and Blackburn 1995); the RNA moiety of telomerase serves as the template for the addition of a characteristic telomere repeat in ciliates (Autexier and Greider 1994) and, despite being absent from the circular genome of prokaryotes, has been considered a candidate for the RNA world (Weiner 1988). The highly conserved structures of the telomerase RNAs of ciliates (Romero and Blackburn 1991; Lingner et al. 1994; Bhattacharyya and Blackburn 1994), as well as regions of absolutely conserved primary sequence essential for the activity of *Tetrahymena* telomerase (Yu et al. 1990), suggest that the RNA has a catalytic activity in telomerase. The RNA alone (without the protein components) forms a tertiary structure similar to that proposed for the RNP on the basis of previous modeling (Romero and Blackburn 1991; Lingner et al. 1994). In addition its helix III may adopt several alternate pseudoknot tertiary structures (Bhattacharyya and Blackburn 1994) with similarities to spliced leader RNAs of *Leptomonas* that may be coupled to an RNA helicase function (Bhattacharyya and Blackburn 1994). Recent results showing that specific mutations in *Tetrahymena* affect fidelity and product dissociation support predictions that at least part of the telomerase active site is composed of RNA (Gilley et al. 1995; Gilley and Blackburn 1996). Moreover, mutations in the RNA component of telomerase causes telomere elongation, suggesting that regulation is affected by such mutations (McEachern and Blackburn 1995).

The protein component of telomerase has interesting features indicative of an ancient origin; the RNP can act as an RNA-dependent RNA polymerase and contains a class II tRNA synthetase motif (Collins et al. 1995), which is interesting in the light of predictions that tRNA-like structures were the first telomeres (Maizels and Weiner 1993, 1994). Weiner (1988) has pointed out that telomerase and the enzyme that adds CCA to the 3′ ends of tRNAs are similar in that they are both primer-dependent, template-independent terminal transferases. The evidence at present supports its existence at least in the RNP world (post translation). Despite recent advances (Lingner et al. 1995; Gilley et al. 1995; Gilley and Blackburn 1996) the evidence for the involvement of the RNA in catalytic activity is not as strong as for the spliceosome or the ribosome. Showing that the RNA of telomerase is directly involved in the synthesis of DNA would, under our criteria, support its existence prior to the origin of translation. This would imply that ancient (RNA) genomes were linear genomes and supports models such as the genomic tag hypothesis of Maizels and Weiner (1993, 1994).

Vault RNAs

The discovery of vault RNPs (Kedersha and Rome 1986) which adhere to the surface of the nuclear membrane and associate with the nuclear pore complex (Chugani et al. 1993) raises difficult questions concerning compartmentation in the ribo-organism. The nature and function of the vRNA moiety are unclear (Nigg 1997), so its full significance for the RNA world remains unknown. It may simply be a form of tag (Maizels and Weiner 1994) for transport to and from the nucleus. The RNA sequence is conserved (Kickhoefer et al. 1993) and if the vRNA turns out to have a role in catalysis or nuclear transport then it may be a relic of an RNA world. These possibilities prompt the consideration of two hypotheses as to the origin of the nucleus:

1. That vault-RNA activity arose de novo after the advent of proteins, which means that the assumptions leading to our model of the RNA world must be reassessed. This option, as does the next, has nuclei arising endogenously.
2. That a pre-nucleus and pre-plasm existed in the RNA world as an aid to separation of replication and transcription (gene expression). Whether such a separation existed is an important question for any RNA

world model, the genomic tag hypothesis is one such proposition (Maizels and Weiner 1993). RNA viruses can separate the two functions and the replication/transcription division is discussed again later. We use the terms pre-nucleus and pre-plasm because membranes are clearly not necessary for compartmentalization in a cell; the nucleolar region within the nucleus is one such example. Even an external membrane in the RNA world has been questioned (Gibson and Lamond 1990).

3. If vRNA turns out to be related to srpRNA it may be derived from the srpRNA of an endosymbiont that formed the nucleus.

The original (and indeed the current) role of vRNAs is an enigma and developments in the field of nuclear transport, especially with regard to the role of vault RNA, are urgently required.

Summary of RNA Processing in the RNA World

Upon examining the use of RNA in modern cells, we observe a general pattern of RNA processing (pre-rRNA, pre-tRNA, and pre-mRNA) where transcripts are processed by an RNP enzyme to yield mature RNA (Fig. 2). Almost all ubiquitous RNAs are involved in processing or assembly of the translation apparatus with telomerase, srpRNA, and vault RNAs as possible exceptions. By our criteria, the general pattern is good evidence that processing is a relic that dates back to the RNA world. The extensive use of RNA in translation shows us an afterimage of an ''evolutionary transition'' (see Maynard Smith and Szathmáry 1995)—how the RNA world brought about translation and began the protein world. There may have been other cascades of RNA processing in the last ribo-organisms but we do not see the products because they have been replaced by proteins in modern organisms. There does not seem to be a general increase in RNP-catalyzed processes during evolution, in agreement with the takeover of catalysis by proteins.

Given our evocation of tRNA in the RNA world, the processing of pre-tRNA transcripts by RNase P is expected to be an ancient RNA world pathway. The pattern is similar to the processing of pre-rRNA by small nucleolar RNAs. Many of the essential functions of ribosomes have been shown to be RNA mediated (Noller 1993a), and there is little doubt about the antiquity of the ribosome. It follows then that the processing of pre-rRNA and pre-tRNA transcripts must also predate translation. Although the rRNA operons of the three ur kingdoms have different gene orders (Pace and Burgin 1990) the cotranscription of rRNA genes is conserved and conspicuous, particularly in eukaryotes where cotranscription is not the norm. The use of RNase P in prokaryotic processing of pre-rRNA- and pre-tRNA-containing transcripts also suggests that this process preceded translation. This processing requires that transcripts and the genome were separate entities in *Riborgis eigensis.*

It is an important conclusion that the last ribo-organism maintained copies of its RNA genome separate from its active ribozymes and that this genome contained sequences additional to those required to code for ribozymes. Thus transcription was even then distinct from replication (Watson et al. 1987). Before giving an overview of the genome of the last ribo-organism we should consider other aspects of metabolism.

Metabolic Complexity

Evidence is weak or lacking for most synthetic reactions and for a plausible energy source (Table 3) in the last ribo-organism. The main evidence comes from the suggestion that ribonucleotide cofactors are relics of an earlier pre-protein metabolism (White 1976, 1982; see Fig. 2) so most classes of biochemical reactions were potentially carried out by RNA enzymes. Benner and Ellington (1987) suggest that it is relatively easy to get ribozymes to catalyze most metabolic steps. Metal ions as cofactors probably also have their origins in the RNA world (Yarus 1993); certainly ribozymes are dependent on metal ions for stability as well as for catalysis (Pyle 1993). Current use of metal ions by proteins reveals that most classes of biochemical reactions can be performed by metalloproteins. Robertson and Miller (1995a) suggest, based on their work on syntheses of amino acid derivatives of uracil, that many of the functions available to proteins were also available to RNA world organisms. The use of either nucleotide cofactors or metal ions in a modern reaction certainly does not automatically place it in the RNA world, though such reactions help demonstrate the potential for catalytic diversity and support the inferred complexity of the RNA world (Table 3). It must be assumed that now-extinct ribozymes with covalently bound cofactor molecules gradually lost the RNA component with the advent of proteins (White 1982). It has been shown that ribozymes can bind cofactors in active sites and can catalyze covalent attachments of NAD^+ and CoA-SH (Breaker and Joyce 1995). Whatever the case, cofactor biochemistry is a major component of extant biochemistry, and this infers that the potential for complexity in the RNA world was indeed quite large. Whether this potential complexity was used to the same extent as at present is unresolved.

Suggestions for metabolic activities that a complex ribo-organism may have carried out are shown in Table 3. As with many situations where fossil evidence is lacking, we cannot be sure of the exactness of the model, only the plausibility (Eigen 1992; Wächtershäuser 1992). The recent surge of in vitro RNA evolution studies (e.g., Doudna et al. 1991; Bartel and Szostak 1993; Lorsch and Szostak 1994; Wilson and Szostak 1995; Illangasekare et al. 1995) bolsters the proposal that RNA is capable of performing many of the catalytic roles currently filled by

Table 3. Other ribozyme-catalyzed functions a complex ribo-organism may have had

Function	Selective advantage	Demonstrated in vitro?	Estimated size (bp)	References
RNA replicase/ polymerase	Necessary for genome replication	Yes	~200 minimum	Doudna et al. (1991) Bartel and Szostak (1993)
Charging of tRNA	Required for an all-RNA ribosome	Yes	~100 (×10)	Illangasekare et al. (1995)
ATPase	Use as ubiquitous energy source	(Yes)	~100	Lorsch and Szostak (1994)
Recombinase	Overcomes Eigen limit and Muller's ratchet; allows for longer chromosomes to be maintained	Spliceosome can carry out cleavage and ligation in *trans*	See table 2 (spliceosome)	Chiara and Reed (1995); Ares and Weiser (1995)
Energy source—via metabolic pathway	Life requires energy; ubiquity of TCA cycle implies inclusion in universal ancestor, and its predecessors	No	?	Wächtershäuser (1990)
Ribose synthesis	High requirements in a ribo-organism, scarce in prebiotic environments	No	?	
Base syntheses	High requirements in a ribo-organism, competition with other nonsynthesizing ribo-organisms	No	?	
Membrane component synthesis	Ability to sequester ions, etc. (see text)	No	?	
Cofactor syntheses	Expands chemistry of reactions	No	?	

proteins. This approach is at present the most fruitful way to examine the extent to which RNA is capable of performing catalytic functions deemed essential in the RNA world.

Physical containment by membranes is an essential part of the viability of extant life, and some form of containment is also deemed necessary in models of the earliest chemical cycles (Wächtershäuser 1990); computer simulations showed hypercycle models were difficult (or impossible) if components were not confined (Niesert et al. 1981). Given the apparent complexity of the RNA world, separation of metabolic processes from the environment is essential; otherwise components at low concentration would never meet, as exemplified by the high concentrations of Mg^{2+} required for ribozyme function (Guerrier-Takada et al. 1983). In contrast to a requirement for compartmentation, Koch (1984) shows that fixation of advantageous mutant genes and the elimination of deleterious ones are retarded with multiple copies of genes (up to about 200) in one ''individual'' or compartment. This indicates that cell division would have been selected for early in the evolution of life. The two possible strategies that could have overcome such concentration dependence of hypercycles are the immobilization of substrates and enzyme cofactors in complexes and compartmentalization via membranes.

In protein enzymes the adenosyl moiety of the nucleotide cofactor binds to the protein, and the cofactor can thus partake in the reaction (Hartman 1975; Eklund et al. 1984). This implies that from one general RNA scaffold, a number of metabolic reactions could be recruited by interchanging the cofactor bound to the RNA (Connell and Christian 1993). Alternatively, a variant of the genomic tag hypothesis (Maizels and Weiner 1993, 1994) could have the first substrate bound to a tRNA-like molecule (much like some of the reactions discussed earlier involving tRNA-bound small molecules). In either case specialization and complexity could be built very simply and quickly by gene duplication. These approaches would enable access to the full range of cofactor-mediated reactions without the substrate diffusing away. Expansion of pathways would rely on the use of the preexisting binding domains and the availability of the appropriate cofactor.

The biochemical nature of membranes in *R. eigensis* is unclear. A terpenoid membrane was included in the palimpsest model (Benner et al. 1989), and recently Ourisson and Nakatani (1994) suggested that the earliest membranes were composed of acyclic polyprenols, though it is currently impossible to reconcile the timeline of events they propose with those for RNA evolution. Overall, compartmentation of some form is essential, but at present we cannot detect any relic. In vitro RNA evolution studies currently offer the only means by which to

address this problem: if RNA enzymes can be evolved that are capable of producing simple membrane components, we can at least get an idea as to the potential complexity of the RNA world membrane.

The Genome of the Last Ribo-organism

The minimum size of the genome must equal or exceed the combined lengths of tRNA, rRNA, the processing molecules (small nuclear and nucleolar RNAs and RNase P), and the hypothetical ribozymes involved in the additional biochemical functions predicted for a complex ribo-organism (Tables 2 and 3). This produces a genome of at least 10,000–15,000 bp. As discussed by Eigen and Schuster (1979), the fidelity of replication places a critical limit on the amount of information that can be maintained by a self-replicating system. Viral RNA replicases and DNA replicases without proofreading functions fall within the range of 10^{-3}–10^{-5} mutations per nucleotide (Domingo and Holland 1994). Even with the use of modern RNA polymerases made of protein, this equates to an Eigen limit on RNA viral genome size of 30 kb. These limits pose critical problems for the maintenance of the large genome in the RNA world inferred from our study and others (White 1976; Benner et al. 1989). Thus, the combined problem of larger genome size, limited by accuracy of replication, is expected to be a major problem at the final stage of the RNA world.

One solution has been to suggest the existence of DNA in the RNA world (Benner et al. 1989), but there is really no evidence for such a proposal. Ribozymes have been retained for RNA processing, and if DNA had existed in the pre-protein world we might have expected ribozymes to still be involved in processing DNA transcripts. Chemical difficulties involved in DNA synthesis, such as the reduction of ribonucleotides to deoxyribonucleotides by a reductase that uses a protein radical to carry out the reaction, make it seem unlikely that such a function could have pre-dated catalytic protein chemistry as RNA is not known to control free radicals (Reichard 1993). The DNA replication apparatus appears too complex for ribozymes (Reanney 1987; Reichard 1993, 1995; Stubbe 1994), which would make it likely that DNA was not used until a reasonable degree of protein catalytic power had evolved. RNA might not be capable of tasks such as the unwinding of the two strands of a DNA double helix, maintenance of supercoiling, and keeping the DNA in an appropriate conformation for the simultaneous replication of leading and lagging strands. DNA may thus be a late addition to the cellular architecture of RNP organisms, and we will work with the assumption that the last ribo-organism still had an RNA genome. The replacement of ribose and uracil to produce a DNA genome hence occurred later. Since the HIV reverse transcriptase can use either RNA or DNA as a template (Peliksa and Benkovic 1992), and *Tetrahymena* telomerase can incorporate RNA instead of DNA (Collins et al. 1995), it seems likely that DNA/RNA substrate/product specificity can evolve relatively easily in proteins.

The last ribo-organism may have had a dsRNA genome; it arises easily by nonseparation of the replication products of a ssRNA genome. Without a nucleus and where cleavage of pre-RNA transcripts occurs, an ssRNA genome would have difficulty maintaining itself when ribozymes are constantly being cleaved from nascent ssRNA transcripts (see earlier discussions). The separation of genotype and phenotype would be more easily accomplished with dsRNA than a nucleus and associated transport mechanisms. One problem is that double-stranded RNA has been shown to slow protein RNA replicases more so than the hairpin loops present in folded RNA intramolecularly (Priano et al. 1987). A dsRNA genome would be more stable (less susceptible to cleavage) than would ssRNAs that are folded into active ribozymes. Furthermore, the maintenance of a genome copy (that is, dsRNA ''gene'' form) of any ribozyme effectively prevents its expression until such time as it is expressed in the ssRNA ''enzyme'' form. Certainly from the point of view of metabolic regulation, this distinction might be important. Whatever the genetic structure of the last ribo-organism, the three steps required in the transition from ssRNA to dsDNA (single to double stranded, ribose to deoxyribose, and uracil to thymine) most likely occurred at different times. Each would presumably allow an increase in accuracy of information storage and copying, allowing a larger genome to be maintained within the Eigen limit.

The complexity of the ribo-organism we describe includes many genes, and separate processes for replication and transcription (though not necessarily a separate apparatus for each). That the last ribo-organism must be able to assemble ribozyme complexes capable of translation (protoribosomes, Poole et al. 1997), complete with aminoacylated tRNAs, is an indication of the complexity of its genome. There are several ways in which such a genome may have been regulated including genomic tags that allow replication to be separated from transcription (Maizels and Weiner 1993) and internal tRNA-like cruciforms in double-stranded RNA which could act as promoters. We can think of pre-tRNA transcripts as ''ribozymogens'' in that they are active only once cleaved. Such ribozymogens and sequence-specific hammerhead-like ribozymes would produce cascades of activation.

Simulations of the selective advantages of joining genes that are components of a hypercycle (Eigen and Schuster 1979) show that once linked genes have arisen, the advantages of keeping mutually beneficial genes together in one replicative unit outweigh the disadvantages of slower replication (Maynard Smith and Szathmáry 1993). Given this, and the universal presence of multi-

gene transcripts for rRNA, the last ribo-organism is expected to have linked genes, though the longer sequence length would lower the effective replication fidelity. One solution to an error catastrophe is the maintenance of multiple copies of the genome, which makes it possible to maintain and store information in spite of low fidelity (Koch 1984). Without a specialized chromosome segregation system, genetic material will tend to increase in any case. In this manner, information stability is conferred through redundancy. For up to 200 copies, polyploidy ensures against loss of genetic information between generations without sacrificing the propensity to fix new mutants within the population (and therefore evolve) (Koch 1984).

As the genome and, concurrently, chromosomes become longer, further measures to maintain genome fidelity are necessary because every unit (chromosome) is prone to picking up sequence errors. Recombination within and between chromosomes is important in avoiding degeneration of genetic information resulting from error accumulation (Muller's ratchet, see Maynard Smith 1978). Recombination offers a means by which to overcome the Eigen limit and maintain relatively large chromosomes; it can be viewed as part of an early gene repair system required even at early stages in the evolution of genomes (Penny 1985). Reanney (1987) has noted that recombination could be achieved simply, and concurrent with RNA replication, by template switching during RNA synthesis. An alternative would be to utilize the cleavage and ligation activities of the spliceosome, via *trans*-splicing (Reanney 1984). In an ancient context the spliceosome could be involved in both the splicing out of introns to produce functional ribozymes, as well as providing a cleavage-ligation function in *trans* to allow limited recombination. A protospliceosome is expected only to be able to act on single-stranded RNA, as is the case with splicing in eukaryotes. Hence recombination between single strands of RNA would precede replication of the second strand. Given the flexibility of RNA enzymes (Kikuchi et al. 1993; Guerrier-Takada et al. 1989) it seems that such a protospliceosome might well have displayed both activities. Combining all these points, including the need to avoid Muller's ratchet, we assume a polyploid fragmented genome of linear chromosomes capable of some recombination (Fig. 4).

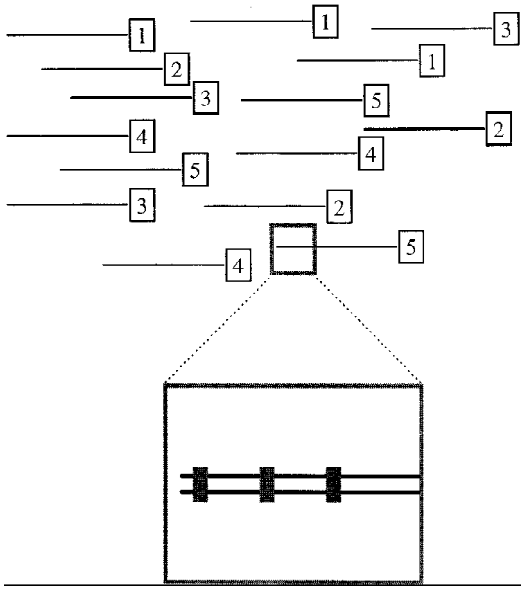

Fig. 4. The likely genome structure of the last ribo-organism. The model we describe for the genome of the last ribo-organism is a linear double-stranded RNA genome fragmented into a number of physically distinct ''chromosomes.'' The genome is present in multiple copies—that is, it is polyploid. (Here it is arbitrarily shown as triploid with five ''chromosomes,'' though there may be anywhere up to 200 copies of the genome—see text.) Polyploidy is presumed to act as a buffer for the low level of replicative fidelity imagined for an organism which utilizes a ribozyme RNA replicase. Additionally, each ''chromosome'' has a number of origins of replication along its length (shown as *solid gray boxes* in the enlarged portion). Gene structure is proposed to be noncontinuous, with intervening sequences separating the coding regions, both within and between genes (not shown). The model does not consider cytological information such as cellular or acellular, membranes present or absent, or details of cellular compartmentation (for instance, a nuclear structure).

Summary and Discussion

This analysis has shown that, given the reasonable assumption of the transfer of catalytic function from RNA → RNP → protein, we can estimate many of the properties of the RNA world just before encoded protein synthesis began. The ancestors of the ribosome, snoRNAs, RNase P, and tRNAs, were a part of the complexity, as were the snRNAs of the spliceosome, and telomerase RNA. At least 10,000–15,000 bp would be involved in these processes. It is likely that ancestors of the signal recognition particle RNA, vault RNA, and RNA editing via base conversion were present, although the specific context in which vault and SRP RNAs were used may well have been different from the modern context. The ribo-organism also made use of pyridine nucleotide cofactors, ribozyme catalysts, and perhaps peptide-stabilized ribozymes for a metabolism that necessarily included many synthetic reactions. Our model for this RNA world is an organism that probably used a double-stranded linear RNA genome that was separate from its single-stranded transcripts—the metabolically active forms. The genome was fragmented into several ''chromosomes'' which were present in multiple copies per cell, each containing genes that were spliced post-transcriptionally to produce functional RNA products.

Although some parts of our model may stand or fall with new experimental evidence, our description of the final complexity of RNA life forms is consistent with known data and has considerable explanatory power. The description of the genome organization and the process-

ing of RNA transcripts provides an explanation for the origin of RNA splicing and the use of RNA in the spliceosome. We favor a historical rather than a functional explanation for the origin of introns (Poole et al. 1997). The origin of snoRNAs and the extensive use of RNA in the ribosome become clear when we realize that translation and ribosome manufacture must have originated in the RNA world.

One of the most crucial aspects of this work is its potential predictive power; it can be used as a basis from which to address specific questions relating to later stages in the evolution of life. For instance, elsewhere (Poole et al. 1997), we describe models for the origin of some of the first introns as well as some of the first genetically encoded proteins, and, moreover, identify putative examples of both. We also use this model to address the sequential evolution of the ribosome. Another question of particular relevance is that of the metabolic and genetic nature of the Last Universal Common Ancestor (LUCA) of all extant life. As the model we have described here is based on metabolic fossils found in extant organisms, these relics must also have existed in the Last Universal Common Ancestor. Hence we have used the model described here as an outgroup to root the tree of life (Poole et al. 1997). One final point is that the overall model we present is far more consistent with a ribo-organism → LUCA → eukaryote-like genome → prokaryote genome series than it is for a ribo-organism → LUCA → prokaryote → eukaryote series.

Note Added at Proof

The role of small nucleolar (sno)RNAs has been expanded still further by the demonstration that they mark uridines to be modified to pseudouridines (Ni et al. 1997; Ganot et al. 1997). This further expands the metabolism of the RNA world, supporting the conclusions of the present paper.

Acknowledgments. We thank Patrick Forterre and an anonymous referee for valuable and constructive comments on the manuscript.

References

Albery WJ, Knowles JR (1976) Evolution of enzyme function and the development of enzyme efficiency. Biochemistry 15:5631–5640

Andersson DI, Hughes D (1996) Muller's ratchet decreases fitness of a DNA-based microbe. Proc Natl Acad Sci USA 93:906–907

Apirion D, Miczak A (1993) RNA processing in prokaryotic cells. Bioessays 15:113–120

Ares M, Weiser B (1995) Rearrangement of snRNA during assembly and function of the spliceosome. Prog Nucleic Acid Res Mol Biol 50:131–159

Autexier C, Greider CW (1994) Functional reconstitution of wild-type and mutant *Tetrahymena* telomerase. Genes Dev 8:563–575

Bachellerie JP, Michot B, Nicoloso M, Balakin A, Ni J, Fournier MJ (1995) Antisense snoRNAs: a family of nucleolar RNAs with long complementarities to rRNA. Trends Biochem Sci 20:261–264

Bartel DP, Szostak JW (1993) Isolation of new ribozymes from a large pool of random sequences. Science 261:1411–1418

Bass BL (1993) RNA editing: new uses for old players in the RNA world. In: Gesteland RF, Atkins JF (eds) The RNA world. Cold Spring Harbor Laboratory Press, New York, pp 383–418

Baumann P, Qureshi SA, Jackson SP (1995) Transcription: new insights from studies on archaea. Trends Genet 11:279–283

Beltrame M, Tollervey D (1992) Identification and functional analysis of two U3 binding sites on yeast pre-ribosomal RNA. EMBO J 11:1531–1542

Benner SA, Ellington AD (1987) The last ribo-organism. Nature 329: 295–296

Benner SA, Ellington AD, Tauer A (1989) Modern metabolism as a palimpsest of the RNA world. Proc Natl Acad Sci USA 86:7054–7058

Bhattacharyya A, Blackburn EH (1994) Architecture of telomerase RNA. EMBO J 13:5721–5731

Blackburn EH (1992) Telomerases. Annu Rev Biochem 61:113–129

Böhler C, Nielsen PE, Orgel LE (1995) Template switching between PNA and RNA oligonucleotides. Nature 376:578–581

Breaker RR, Joyce GF (1995) Self-incorporation of coenzymes by ribozymes. J Mol Evol 40:551–558

Brown JW, Pace NR (1992) Ribonuclease P RNA and protein subunits from bacteria. Nucleic Acids Res 20:1451–1456

Cavaille J, Nicoloso M, Bachellerie J-P (1996) Targeted ribose methylation of RNA in vivo directed by tailored antisense RNA guides. Nature 383:732–735

Cech TR (1986) The generality of self-splicing RNA: relationship to nuclear mRNA splicing. Cell 44:207–210

Cech TR, Uhlenbeck OC (1994) Hammerhead nailed down. Nature 372:39–40

Chiara MD, Reed R (1995) A two-step mechanism for 5′ and 3′ splice-site pairing. Nature 375:510–513

Chugani DC, Rome LH, Kedersha NL (1993) Evidence that vault ribonucleoprotein particles localize to the nuclear pore complex. J Cell Sci 106:23–29

Coetzee T, Herschlag D, Belfort M (1994) *Escherichia coli* proteins, including ribosomal protein S12, facilitate in vitro splicing of phage T4 introns by acting as RNA chaperones. Genes Dev 8:1575–1588

Cohn M, Blackburn EH (1995) Telomerase in yeast. Science 269:396–400

Collins K, Kobayashi R, Greider CW (1995) Purification of *Tetrahymena* telomerase and cloning of genes encoding the two protein components of the enzyme. Cell 81:677–686

Connell GJ, Christian EL (1993) Utilization of cofactors expands metabolism in a new RNA world. Orig Life Evol Biosph 23:291–297

Cooperman BS, Wooten T, Romero DP, Traut RR (1995) Histidine 229 in protein L2 is apparently essential for 5OS peptidyl transferase activity. Biochem Cell Biol 73:1087–1094

Crick FHC (1968) The origin of the genetic code. J Mol Biol 38:367–379

Cruz-Reyes J, Sollner-Webb B (1996) Trypanosome U-deletional RNA editing involves guide RNA-directed endonuclease cleavage, terminal U exonuclease, and RNA ligase activities. Proc Natl Acad Sci USA 93:8901–8906

Curnow AW, Ibba M, Söll D (1996) tRNA-dependent asparagine formation. Nature 382:589–590

Diener TO (1993) Small pathogenic RNAs of plants: living fossils of the RNA world? In: Ponnamperuma C, Chela-Fores J (eds) Chemical evolution: origin of life. A Deepak Publishing, Hampton, VA, pp 69–83

Domingo E, Holland JJ (1994) Mutation rates and rapid evolution of RNA viruses. In: Morse SS (ed) The evolutionary biology of viruses. Raven Press, New York, pp 161–183

Doudna JA, Couture S, Szostak JW (1991) A multisubunit ribozyme

that is a catalyst of and template for complementary strand RNA synthesis. Science 251:1605–1608
Draper DE (1995) Protein-RNA recognition. Annu Rev Biochem 64: 593–620
Eck RV, Dayhoff MO (1966) Evolution of the structure of ferredoxin based on living relics of primitive amino acid sequences. Science 152:363–366
Eigen M (1992) Steps toward life: a perspective on evolution. Oxford University Press, Oxford
Eigen M (1993) The origin of genetic information: viruses as models. Gene 135:37–47
Eigen M, Schuster P (1977) The hypercycle: a principle of natural self-organization. Part A: emergence of the hypercycle. Naturwissenschaften 64:541–565
Eigen M, Schuster P (1979) The hypercycle: a principle of natural self-organization. Springer-Verlag, Berlin
Eigen M, Winkler-Oswatitsch R (1981) Transfer-RNA, an early gene? Naturwissenschaften 68:282–292
Eigen M, Lindemann BF, Tietze M, Winkler-Oswatitsch R, Dress A, von Haeseler A (1989) How old is the genetic code? Statistical geometry of tRNA provides an answer. Science 244:673–679
Eklund H, Samama J-P, Jones TA (1984) Crystallographic investigations of NAD binding to horse liver alcohol dehydrogenase. Biochemistry 23:5982–5996
Ertem G, Ferris JP (1996) Synthesis of RNA oligomers on heterogeneous templates. Science 379:238–240
Forterre P (1996) A hot topic: the origin of hyperthermophiles. Cell 85:789–792
Fung PA, Gaertig J, Gorovsky MA, Hallberg RL (1995) Requirement of a small cytoplasmic RNA for the establishment of thermotolerance. Science 268:1036–1039
Fung PA, Labrecque R, Pederson T (1997) RNA-dependent phosphorylation of a nuclear RNA binding protein. Proc Natl Acad Sci USA 94:1064–1068
Gesteland RF, Atkins JF (eds) (1993) The RNA world. Cold Spring Harbor Laboratory Press, New York
Gibson TJ, Lamond AI (1990) Metabolic complexity of the RNA world and implications for the origin of protein synthesis. J Mol Evol 30:7–15
Gilbert W (1986) The RNA world. Nature 319:618
Gilley D, Blackburn EH (1996) Specific RNA residue interactions required for enzymatic functions of *Tetrahymena* telomerase. Mol Cell Biol 16:66–75
Gilley D, Lee MS, Blackburn EH (1995) Altering specific telomerase RNA template residues affects active site function. Genes Dev 9:2214–2226
Gordon KHJ (1995) Were RNA replication and translation directly coupled in the RNA (+protein?) world? J Theor Biol 173:179–1931
Greider CW, Blackburn EH (1989) A telomeric sequence in the RNA of *Tetrahymena* telomerase required for telomere repeat synthesis. Nature 337:331–337
Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S (1983) The RNA moiety of RNase P is the catalytic subunit of the enzyme. Cell 35:849–857
Guerrier-Takada C, Lumelsky N, Altman S (1989) Specific interactions in RNA enzyme-substrate complexes. Science 246:1578–1584
Gupta RS, Singh B (1994) Phylogenetic analysis of 70kD heat shock protein sequences suggests a chimaeric origin for the eukaryotic nucleus. Curr Biol 4:1104–1114
Hartman H (1975) Speculations on the origin and evolution of metabolism. J Mol Evol 4:359–370
Hasegawa M, Yano T, Miyata T (1984) Evolutionary implications of error amplification in the self-replicating and protein-synthesizing machinery. J Mol Evol 20:77–85
Herschlag D, Cech TR (1990) Catalysis of RNA cleavage by the *Tetrahymena thermophila* ribozyme. 1. Kinetic description of the reaction of an RNA substrate complementary to the active site. Biochemistry 29:10159–10171
Herschlag D, Khosla M, Tsuchihashi Z, Karpel RL (1994) An RNA chaperone activity of non-specific RNA binding proteins in hammerhead ribozyme catalysis. EMBO J 13:2913–2924
Illangasekare M, Sanchez G, Nickles T, Yarus M (1995) Aminoacyl-RNA synthesis catalyzed by an RNA. Science 267:643–647
Jacob F (1977) Evolution and tinkering. Science 196:1161–1166
Jacobsen JR, Schultz PG (1994) Antibody catalysis of peptide bond formation. Proc Natl Acad Sci USA 91:5888–5892
Jeffares DC, Poole AM, Penny D (1995) Pre-rRNA processing and the path from the RNA world. Trends Biochem Sci 20:298–299
Joyce GF, Orgel LE (1993) Prospects for understanding the origin of the RNA world. In: Gesteland RF, Atkins JF (eds) The RNA world. Cold Spring Harbor Laboratory Press, New York, pp 1–25
Kable ML, Seiwert SD, Heidmann S, Stuart K (1996) RNA editing: a mechanism for gRNA-specified uridylate insertion into precursor mRNA. Science 273:1189–1195
Kaine BP (1990) Structure of the archaebacterial 7S RNA molecule. Mol Gen Genet 221:315–321
Kamaishi T, Hashimoto T, Nakamura Y, Nakamura F, Murata S, Okada N, Okamoto K, Shimizu M, Hasegawa M (1996) Protein phylogeny of translation elongation factor EF-1α suggests microsporidians are extremely ancient eukaryotes. J Mol Evol 42:257–263
Kedersha NL, Rome LH (1986) Isolation and characterization of a novel ribonucleoprotein particle: large structures contain a single species of small RNA. J Cell Biol 103:699–709
Keiler KC, Waller PRH, Sauer RT (1996) Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. Science 271:990–993
Kickhoefer VA, Searles RP, Kedersha NL, Garber ME, Johnson DL, Rome LH (1993) Vault ribonucleoprotein particles from rat and bullfrog contain a related small RNA that is transcribed by RNA polymerase III. J Biol Chem 268:7868–7873
Kikuchi Y, Sasaki-Tozawa N, Suzuki K (1993) Artificial self-cleaving molecules consisting of a tRNA precursor and the catalytic RNA of RNase P. Nucleic Acid Res 21:4685–4689
Kiss T, Filipowicz W (1993) Small nucleolar RNAs encoded by introns of the human cell cycle regulatory gene RCCI. EMBO J 12:2913–2920
Kiss T, Filipowicz W (1995) Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns. Genes Dev 9:1411–1424
Kiss-László Z, Henry Y, Bachellerie J-P, Caizergues-Ferrer M, Kiss T (1996) Site-specific ribose methylation of preribosomal RNA; a novel function for small nucleolar RNAs. Cell 85:1077–1088
Koch AL (1984) Evolution vs the number of gene copies per primitive cell. J Mol Evol 20:71–76
LaGrandeur TE, Darr SC, Haas ES, Pace NR (1993) Characterization of the RNase P RNA of *Sulfolobus acidocaldarius.* J Bacteriol 175:5043–5048
Lahav N, White D, Chang S (1978) Peptide formation in the prebiotic era: thermal condensation of glycine in fluctuating clay environments. Science 201:67–69
Lazcano A, Miller SL (1996) The origin and early evolution of life: prebiotic chemistry, the pre-RNA world, and time. Cell 85:793–798
Lesser CF, Guthrie C (1993) Mutations in U6 snRNA that alter splice site specificity: implications for the active site. Science 262:1982–1988
Leverette RD, Andrews MT, Maxwell ES (1992) Mouse U14 snRNA is a processed intron of the cognate hsc70 heat shock pre-messenger RNA. Cell 71:1215–1221
Lewin B (1994) Genes V. Oxford University Press, New York
Lillehaug JR, Kleppe K (1975) Kinetics and specificity of T4 polynucleotide kinase. Biochemistry 14:1221–1225
Lingner J, Hendrick LL, Cech TR (1994) Telomerase RNAs of different ciliates have a common secondary structure and a permuted template. Genes Dev 8:1984–1998
Lingner J, Cooper JP, Cech TR (1995) Telomerase and DNA end

replication: no longer a lagging strand problem? Science 269:1533–1534

Liu F, Altman S (1994) Differential evolution of substrates for an RNA enzyme in the presence and absence of its protein cofactor. Cell 77:1093–1100

Liu J, Maxwell ES (1990) Mouse U14 snRNA is encoded in an intron of the mouse cognate hsc70 heat shock gene. Nucleic Acids Res 18:6565–6571

Lonergan KM, Gray MW (1993) Editing of transfer RNAs in *Acanthamoeba castellanii* mitochondria. Science 259:812–816

Lorsch JR, Szostak JW (1994) In vitro evolution of new ribozymes with polynucleotide kinase activity. Nature 371:31–36

Madhani HD, Guthrie C (1994a) Randomization-selection analysis of snRNAs in vivo: evidence for a tertiary interaction in the spliceosome. Genes Dev 8:1071–1086

Madhani HD, Guthrie C (1994b) Dynamic RNA-RNA interactions in the spliceosome. Annu Rev Genet 28:1–26

Maizels N, Weiner AM (1987) Peptide-specific ribosomes, genomic tags, and the origin of the genetic code. Cold Spring Harb Symp Quant Biol 52:743–749

Maizels N, Weiner AM (1993) The genomic tag hypothesis: modern viruses as molecular fossils of ancient strategies for genomic replication. In: Gesteland RF, Atkins JF (eds) The RNA world. Cold Spring Harbor Laboratory Press, New York, pp 577–602

Maizels N, Weiner AM (1994) Phylogeny from function: evidence from the molecular fossil record that tRNA originated in replication, not translation. Proc Natl Acad Sci USA 91:6729–6734

Maynard Smith J (1978) The evolution of sex. Cambridge University Press, Cambridge

Maynard Smith J, Szathmáry E (1993) The origin of chromosomes I. Selection for linkage. J Theor Biol 164:437–446

Maynard Smith J, Szathmáry E (1995) The major transitions in evolution. W H Freeman, Oxford

Maxwell ES, Fournier MJ (1995) The small nucleolar RNAs. Annu Rev Biochem 35:897–934

McEachern MJ, Blackburn EH (1995) Runaway telomere elongation caused by telomerase RNA gene mutations. Nature 376:403–409

Miller JD, Wilhelm H, Gierasch L, Gilmore R, Walter P (1993) GTP binding and hydrolysis by the signal recognition particle during initation of protein translocation. Nature 366:351–354

Miller JD, Bernstein HD, Walter P (1994) Interaction of *E. coli* Ffh/4.5S RNP and FtsY mimics that of mammalian signal recognition particle and its receptor. Nature 367:657–659

Moazed D, Noller HF (1991) Sites of interaction of the CCA end of peptidyl-tRNA with 23S rRNA. Proc Natl Acad Sci USA 88:3725–3728

Morrissey JP, Tollervey D (1993) Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis. Mol Cell Biol 13:2469–2477

Morrissey JP, Tollervey D (1995) Birth of the snoRNPs: the evolution of RNase MRP and the eukaryotic pre-rRNA-processing system. Trends Biochem Sci 20:78–82

Nicoloso M, Qu LH, Michot B, Bachellerie J-P (1996) Intron-encoded, antisense small nucleolar RNAs—the characterization of nine novel species points to their direct role as guides for the 2-O-methylation of rRNAs. J Mol Biol 260:178–195

Niesert U, Harnasch D, Bresch C (1981) Origin of life between Scylla and Charybdis. J Mol Evol 17:348–353

Nigg EA (1997) Nucleocytoplasmic transport: signals, mechanisms and regulation. Nature 386:779–787

Noller HF (1993a) On the origin of the ribosome: coevolution of subdomains of tRNA and rRNA. In: Gesteland RF, Atkins JF (eds) The RNA world. Cold Spring Harbor Laboratory Press, New York, pp 137–156

Noller HF (1993b) Peptidyl transferase: protein, ribonucleoprotein, or RNA? J Bacteriol 175:5297–5300

Noller HF, Hoffarth V, Zimniak L (1992) Unusual resistance of peptidyl transferase to protein extraction procedures. Science 256:1416–1419

Orgel LE (1963) The maintenance of the accuracy of protein synthesis and its relevance to ageing. Proc Natl Acad Sci USA 49:517–521

Orgel LE (1968) Evolution of the genetic apparatus. J Mol Biol 38:381–393

Ourisson G, Nakatani Y (1994) The terpenoid theory of the origin of cellular life: the evolution of terpenoids to cholesterol. Chem Biol 1:11–23

Pace NR, Burgin AB (1990) Processing and evolution of the rRNAs. In: Hill WE, Dahlberg A, Garrett RA, Moore PB, Schlessinger D, Warner JR (eds) The ribosome: structure, function, evolution. American Society for Microbiology, Washington, pp 417–425

Peculis BA, Steitz JA (1994) Sequence and structural elements critical for U8 function in *Xenopus* oocytes are evolutionally conserved. Genes Dev 8:2241–2255

Peliska JA, Benkovic SJ (1992) Mechanism of DNA strand transfer reactions catalysed by HIV-1 reverse transcriptase. Science 258:1112–1118

Penny D (1985) The evolution of meiosis and sexual reproduction. Biol J Linn Soc 25:209–220

Penny D (1988) What was the first living cell? Nature 331:111–112

Poole AM, Jeffares DC, Penny D (1997) The path from the RNA world. J Mol Evol 45:1–18

Potter S, Durovic P, Dennis PP (1995) Ribosomal RNA precursor processing by a eukaryotic U3 small nucleolar RNA-like molecule in an archeon. Science 268:1056–1060

Priano C, Kramer FR, Mills DR (1987) Evolution of the RNA coliphages: the role of secondary structures during RNA replication. Cold Spring Harb Symp Quant Biol 52:321–330

Purohit P, Stern S (1994) Interactions of a small RNA with antibiotic and RNA ligands of the 30S subunit. Nature 370:659–662

Pyle AM (1993) Ribozymes: a distinct class of metallozymes. Science 261:709–714

Pyle AM, Murphy FL, Cech TR (1992) RNA substrate binding site in the catalytic core of the *Tetrahymena* ribozyme. Nature 358:123–128

Radzicka A, Wolfenden R (1995) A proficient enzyme. Science 267:90–93

Reanney DC (1979) RNA splicing and polynucleotide evolution. Nature 277:598–600

Reanney DC (1984) RNA splicing as an error-screening mechanism. J Theor Biol 110:315–321

Reanney DC (1987) Genetic error and genome design. Cold Spring Harb Symp Quant Biol 52:751–757

Reichard P (1993) From RNA to DNA, why so many ribonucleotide reductases. Science 260:1773–1777

Reichard P (1995) To be there when the picture is painted. Annu Rev Biochem 64:1–28

Retallack DM, Friedman DI (1995) A role for a small stable RNA in modulating the activity of DNA-binding proteins. Cell 83:227–235

Rivera MC, Lake JA (1992) Evidence that eukaryotes and eocyte prokaryotes are immediate relatives. Science 257:74–76

Roberts WSL, Strominger JL, Söll D (1968) Biosynthesis of peptidoglycan of bacterial cell walls. VI Incorporation of L-threonine into interpeptide in *Micrococcus roseus.* J Biol Chem 243:749–746

Robertson MP, Miller SL (1995a) An efficient prebiotic synthesis of cytosine and uracil. Nature 375:772–774

Robertson MP, Miller SL (1995b) Prebiotic synthesis of 5-substituted uracils: a bridge between the RNA world and the DNA-protein world. Science 268:702–705

Rogers KC, Söll D (1995) Divergence of glutamate and glutamine aminoacylation pathways: providing the evolutionary rationale for mischarging. J Mol Evol 40:476–481

Romero DP, Blackburn EH (1991) A conserved secondary structure for telomerase RNA. Cell 67:343–353

Samaha RR, Green R, Noller HF (1995) A base pair between tRNA and

23S rRNA in the peptidyl transferase centre of the ribosome. Nature 377:309–314

Schön A, Krupp G, Gough S, Berry-Lowe S, Kannangara CG, Söll D (1986) The RNA required in the first step of chlorophyll synthesis is a chloroplast glutamate tRNA. Nature 322:281–284

Seiwert SD, Heidmann S, Stuart K (1996) Direct visualization of uridylate deletion in vitro suggests a mechanism for kinetoplastid RNA editing. Cell 84:831–841

Sharp PA (1994) Split genes and RNA splicing. Cell 77:805–815

Sollner-Webb B (1993) Novel intron-encoded small nucleolar RNAs. Cell 75:403–405

Stubbe J (1994) Controlling radical reactions. Nature 370:502

Sullenger BA, Cech TR (1994) Ribozyme-mediated repair of defective mRNA by targeted *trans*-splicing. Nature 371:619–622

Sun J, Manly JL (1995) A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing Genes Dev 9:843–854

Szewczak AA, Moore PB (1995) The sarcin/ricin loop, a modular RNA. J Mol Biol 247:81–98

Tarn W-Y, Steitz JA (1996) Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. Science 273:1824–1832

Tu G-F, Reid GE, Zhang J-G, Moritz RL, Simpson RJ (1995) C-terminal extension of truncated recombinant proteins in *Escherichia coli* with a 10Sa RNA decapeptide. J Biol Chem 270:9322–9326

Tycowski KT, Shu M-D, Steitz JA (1996) A mammalian gene with introns instead of exons generating stable RNA products. Nature 379:464–466

von Ahsen U, Noller HF (1995) Identification of bases in 16S rRNA essential for tRNA binding at the 30S ribosomal P site. Science 267:234–237

Wächtershäuser G (1990) Evolution of the first metabolic cycles. Proc Natl Acad Sci USA 87:200–204

Wächtershäuser G (1992) Groundworks for an evolutionary biochemistry: the Iron-Sulphur World. Prog Biophys Mol Biol 58:85–201

Walbot V (1991) RNA editing fixes problems in plant mitochondrial transcripts. Trends Genet 7:37–39

Watson JD, Hopkins NH, Roberts JW, Steitz JA, Weiner AM (1987) Molecular biology of the gene. 4th ed. Benjamin/Cummings, Menlo Park, CA, pp 1098–1163

Weiner AM (1988) Eukaryotic nuclear telomeres: molecular fossils of the RNP world? Cell 52:155–157

Weiner AM (1993) mRNA splicing and autocatalytic introns: distant cousins or the products of chemical determinism? Cell 72:161–164

White HB (1976) Coenzymes as fossils of an earlier metabolic state. J Mol Evol 7:101–104

White HB (1982) Evolution of coenzymes and the origin of pyridine nucleotides. In: Everse J, Anderson B, You K (eds) The pyridine nucleotide coenzymes. Academic Press, New York, pp 1–17

White HD, Erickson JC (1980) Catalysis of peptide bond formation by histidyl histidine in a fluctuating clay environment. J Mol Evol 16:279–290

Wilson C, Szostak JW (1995) In vitro evolution of a self-alkylating ribozyme. Nature 374:777–782

Wise JA (1993) Guides to the heart of the spliceosome. Science 262: 1978–1979

Yang J, Cedergren R, Nadal-Ginard B (1994) Catalytic activity of an RNA domain derived from the U6-U4 RNA complex. Science 263: 77–81

Yarus M (1993) How many catalytic RNAs? Ions and the Cheshire cat conjecture. FASEB J 7:31–39

Yu G-L, Bradley JD, Attardi LD, Blackburn EH (1990) *In vivo* alteration of telomere sequences and senescence caused by mutated *Tetrahymena* telomerase RNAs. Nature 344:126–132