International Journal of Astrobiology 4 (1): 33–41 (2005) Printed in the United Kingdom
doi:10.1017/S1473550405002521

Relicts and models of the RNA world

Kirsi Lehto and Alexey Karetnikov

Laboratory of Plant Physiology and Molecular Biology, FI-20014 University of Turku, Finland
e-mail: klehto@utu.fi

Abstract: It is widely believed that the current DNA–RNA–protein-based life forms have evolved from preceding RNA–protein-based life forms, and these again, from mere RNA replicons. By rationale, it can be assumed that the early RNA replicons were fully heterotrophic in terms of obtaining all their building blocks from their environment. In the absence of protein catalysts, their essential life functions had to be mediated by simple functional structures and mechanisms, such as RNA secondary structures, RNA–RNA interactions and RNA-mediated catalysis, and possibly by catalytic minerals or clays. The central role of RNA catalysts in early life forms is supported by the fact that several catalytic RNAs still perform central biological functions in current life forms, and at least some of these may be derived as molecular relicts from the early RNA-based life. The RNA-catalysed metabolic reactions and molecular fossils are more conserved in the eukaryotic life forms than in the prokaryotes, suggesting that the linear eukaryote genomes may more closely resemble the structure and function of the early RNA replicons, than what do the circular prokaryote genomes. Present-day RNA viruses and viroids utilize ultimately simple life strategies, which may be similar to those used by the early RNA replicons. Thus, molecular and functional properties of viruses and viroids may be considered as examples or models of the structures and replication mechanisms, which might have been used for the replication of the early biopolymers.

Received 2 February 2005, accepted 20 April 2005

Key words: autocatalysis, molecular evolution, origin of life, polymers, replication, translation, viroids, viruses.

Introduction

All current life forms on the Earth are cellular. They are based on complex biochemistry, coordinated by DNA-encoded genetic information and driven by protein-mediated catalysis. This life strategy was supposedly used already by the progenitor, or the last common ancestor (LCA), of all cellular life forms, and from there on, has been conserved in all descending species (Penny 1988; Woese 1998). It is clear that the DNA-based life, with its complex, protein-based maintenance (replication) and its multi-step flow of genetic information from the DNA to proteins, had to be derived from simpler life forms. The evolution had to initiate from the simplest polymers capable of spontaneous replication, and proceed through multiple stages with gradually increasing complexity and functionality (Orgel 2004). As no pre-existing biochemical catalysts were available, the replication of the earliest polymers had to be catalysed either by themselves, or possibly by inorganic catalytic templates, such as clays (James & Ellington 1998; Ferris 2002) or metal sulphides (Wächtershäuser 1998; Martin & Russell 2002) or mediated by physical processes, such as concentration-driven polymerization (Lathe 2005).

It is commonly believed that the DNA-based genetic information was preceded by RNA-encoded genetic information, which would have formed a RNA–protein-based life strategy (Maizels & Weiner 1994; Poole *et al*. 1998; Meli *et al*. 2001; Orgel 2004), although this view has been debated (Dworkin *et al*. 2003). Prior to the invention of the genetic code and the protein synthesis machinery, the RNA–protein-based life forms must have been preceded by protein-independent replicons, presumably RNA molecules, which later evolved into the central components of the protein-coding and synthesis machinery (Fig. 1). This essential precursor stage of evolution, preceding the protein-catalysed life forms, has been called the RNA world (Gilbert 1986). It is believed that many essential functions in the RNA world were mediated by catalytic RNA sequences. With the evolution of the DNA-based genomes and protein-mediated catalysis, the original RNA-based life forms disappeared. However, some central functions in present-day cells are still catalysed by RNA, or utilize co-enzymes with ribonucleotide co-factors, and it is believed that these functional RNAs or nucleotide moieties may be molecular relicts or fossils of the pre-cellular ancestor (Maizels & Weiner 1994; Jeffares *et al*. 1998; Meli *et al*. 2001; Orgel 2004).

Within cellular life forms there also exists a separate domain of replicating molecules, i.e. viruses and viroids. They are fully dependent on their hosts, and have survived through long-term evolution within their suitable host species to produce the currently existing molecular parasites (Roossinck 1997; Hull 2002). The origin of viruses or viroids is not known: They may originate from the RNA world, and predate their DNA-based cellular hosts, or they may be derivatives of

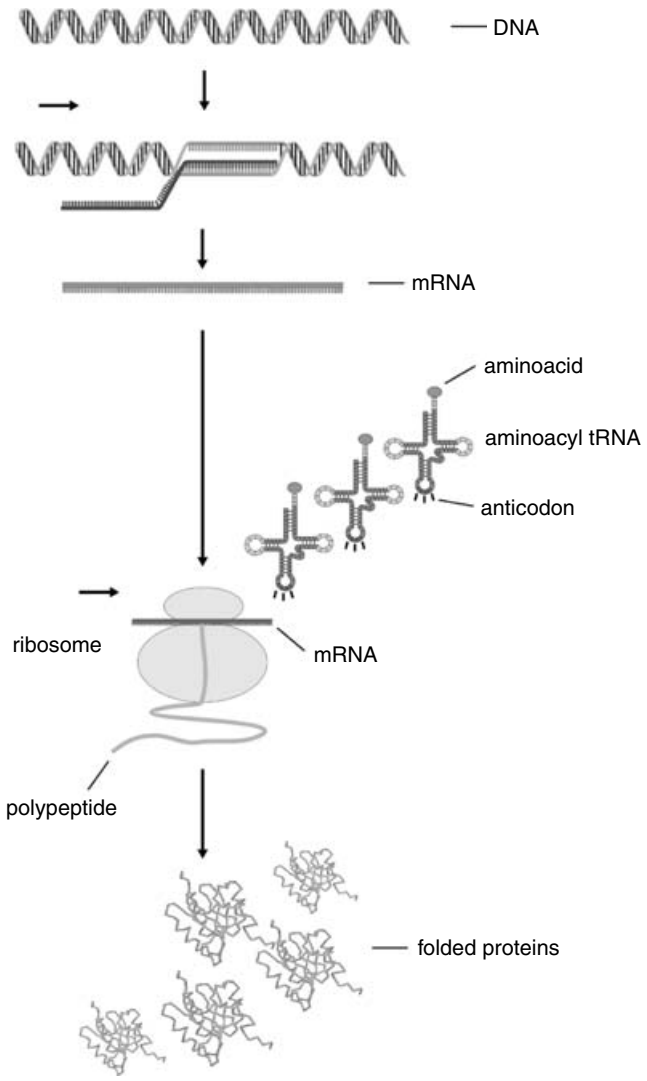

Fig. 1. Flowchart of the expression of genomic information. The nucleotide sequence of one gene is copied (transcribed) from DNA into messenger RNA (mRNA). The mRNA associates with the two subunits of the ribosome, which advances along the mRNA and identifies the genetic code, as nucleotides triplets, from the mRNA. The nucleotide triplets are recognized by the anticodon triplets of the aminoacylated transfer RNAs, which enter into the ribosome complex and bring in the specific amino acids corresponding to each triplet codon (Kapp & Lorsch 2004). The catalytic RNA component of the ribosome forms the peptide bond between the incoming and the adjacent amino acid (Steitz & Moore 2003; Steitz 2005). The produced amino acid chain is called a polypeptide and this folds into the three-dimensional structure of a mature, functional protein. The eukaryotic ribosome complex contains a total of four different RNA molecules and 79 (for mammals and fungi) or 74 (for plants) different protein components (Veuthey & Bittar 1998; Nakao *et al*. 2004; Chang *et al*. 2005). Many other proteins are required in each step of the gene expression, i.e. in the transcription, in the processing of the mRNA (in eukaryotes) and in translation initiation, elongation and termination.

the DNA-based life forms (Maizels & Weiner 1994; Robertson & Neel 1999; Semancik & Duran-Vila 1999). In spite of the uncertainty of their initial origins, they seem to bear interesting functional and structural similarity to the primitive RNA replicons, and therefore they can be considered as functional models, or maybe even functional relicts of the RNA world (Maizels & Weiner 1994). In this paper we review some examples of the molecular relicts and functional models of the RNA world, which exist in contemporary cellular life and in viruses and viroids. We also discuss the putative evolutionary routes that may have been used to convert the molecular replicons to central cellular structures and functions.

Origins of the RNA world

The early existence of a RNA world is strongly supported as a necessary precursory stage in the evolution of the DNA–protein world, but the origin of the functional RNA world itself is not easily comprehensible. According to the RNA-world hypothesis, the very first RNA polymers were spontaneously assembled from pre-existing ribonucleotides, and from some point on, acquired the potential for spontaneous amplification via autocatalytic replication. This precludes prior, non-enzymatic synthesis of the ribonucleotides (in significant amounts), and this again precludes the non-enzymatic synthesis of their components, i.e. of the ribose sugar, and of purine and pyrimidine bases. Also, for assembly of present-day-like nucleosides, the bases had to be combined with β-D-ribose sugars, and be activated at the 5′ carbon with phosphate moieties (or with some other activating compound). The possible chemical routes leading to the prebiotic, or non-enzymatic synthesis of these precursors have been intensively studied and extensively reviewed (e.g. Miller 1998; Schwartz 1998; Orgel 2004). According to this literature, the synthesis of purine bases can be explained via fairly simple prebiotic reactions starting from hydrogen cyanide, possibly including formamide; pyrimidine bases can be formed starting from cyanoacetylene, or from cyanoacetaldehyde and urea (or, further back, from methane and molecular nitrogen). The yields of these synthesis reactions are very low, and they seem to proceed best in eutectic solutions, i.e. in solutions concentrated by freezing (Orgel 2004). This point is in strong contrast to the other hypothesis suggesting that the most plausible site for the synthesis of the reduced precursors of the first biomolecules would have been in hot, high-energy conditions, such as hydrothermal vents (Martin & Russell 2002).

Ribose sugar can be formed, but in very low yields, in the so-called formose reaction starting from formaldehyde. The reaction produces a wide variety of different pentose and hexose sugars, and enrichment of ribose from the mixture is very problematic, as is the instability of ribose – although recent results show that these problems may be helped by Pb-catalysis of the formose reaction, and by calcium borate-mediated stabilization of ribose (Zubay & Mui 2001; Ricardo *et al*. 2004). Furthermore, the purine and pyrimidine bases need to be covalently bound, in β-orientation, with the 1′ carbon of ribose to form nucleosides. This reaction can be activated by heating, and produces purine nucleosides,

https://doi.org/10.1017/S1473550405002521 Published online by Cambridge University Press

although with very low yield. Routes for the prebiotic synthesis of pyrimidine nucleosides are not well known, but they may be feasible via multiple sugar phosphate intermediates (Zubay & Mui 2001). Also, the prebiotic chemical activation of nucleosides is problematic, because soluble phosphate molecules would have not been readily available in the prebiotic world. However, phosphorylation of nucleosides could have happened in dilute solutions of calcium phosphate (hydroxylapatite), in the presence of urea, ammonium chloride and with heating. Also volcanically produced linear polyphosphates, or their breakdown products, could have functioned as prebiotic phosphorylation agents (Orgel 2004 and references therein).

Prebiotic synthesis of nucleotides would have yielded a racemic mixture of a variety of different nucleotide analogues, both in α and in β, and in L and D isoforms. Spontaneous polymerization reactions of such (randomly) phosphorylated nucleotides would have been very slow, and would have led to a wide variety of different linkages, formed between a whole variety of different nucleotide analogues and isoforms. However, most of such randomly linked oligopolymers or polymers would not have been extendable, or functional as templates for replication, but only those formed by 5′–3′ phosphodiester linkages between β-D-nucleotides would have been functional (Joyce 2002; Orgel 2004). The polymerization reaction *per se* could have been essentially enhanced by catalysis by metal ions, and preferably, happened in eutectic solutions (Kanavarioti *et al.* 2001; Monnard *et al.* 2003).

Altogether, prebiotic formation of RNA nucleotides and their polymers, starting from small molecular precursors (such as HCN, NH_3, CH_2O, CH_4, PO_4^{3-}, H_2O) involves so many unlikely or adverse chemical steps, that it has been postulated that the RNA synthesis had to be preceded by simpler chemical polymers. Polymers proposed as hypothetical predecessors of the RNA world include, for example, peptide nucleic acids (PNA), threose nucleic acids (TNA), and glycerol- and pyranosyl-derived nucleic acids (Miller 1998; Schwartz 1998; Joyce 2002; Orgel 2004). Indeed, information transfer from PNA to RNA has been experimentally demonstrated (Schmidt *et al.* 1997). Some other interesting hypotheses have also been proposed, for instance, a hypothetical proto-RNA quadruplex consisting of four polymer strands, the synthesis of which could have been assisted by a so-called 'molecular midwife', a flat hydrophobic molecule that might be similar to phthalocyanine, no longer existing on Earth (Hud & Anet 2000). It is also postulated that instead of organic polymers, entirely different molecules, such as minerals, may have contained the earliest pre-genetic (replicating) information (Cairns-Smith 1982). Very little is known, as yet, concerning the available starting materials, the conditions under which the prebiotic chemistry took place, and what chemical and physical reactions and selection processes were driving the increasing complexity of the early polymers. Thus, the stages preceding the RNA world remain enigmatic, and we do not attempt to dwell on them any further. However, during the molecular evolution certain functional and heritable properties had to be acquired, gradually, and we aim to discuss these as the essential features of the postulated RNA world.

Catalysts of the RNA world and their relicts in present-day life forms

One support for the hypothesis of the early RNA replicons is the central role of different RNA molecules in present-day cellular life (Fig. 1): genetically encoded proteins can be synthesized only from RNA templates, as the DNA-based genetic information is made functional only via copying (transcription) into messenger RNA (mRNA). Transfer RNAs and ribosomal RNAs (tRNAs and rRNAs, respectively) are needed as the functional components of the protein synthesis (translation) machinery, and indeed, 23S rRNA catalyses peptide bond formation inside a large subunit of ribosomes (Zhang & Cech 1997; Steitz & Moore 2003; Steitz 2005).

As no protein enzymes were available to mediate the essential functions of the early replicons, it is postulated that many such functions were catalysed by the replicating RNAs themselves. The repertoire of RNA-catalysed reactions, demonstrated either *in vivo* or *in vitro*, includes RNA polymerization, amino acid activation, aminoacylation of tRNAs, cutting and ligation of RNAs, carbon–carbon bond formation and peptide bond formation (Zhang & Cech 1997; James & Ellington 1998; Lee *et al.* 2000; Jäschke 2001; Johnston *et al.* 2001; Kumar & Yarus 2001; Doudna & Cech 2002; Joyce 2002; Steitz & Moore 2003; Vlassov *et al.* 2004; Li & Huang 2005; Steitz 2005). Also, limited spontaneous polymerization of ribonucleotides can take place in icy environments, aided by catalytic metal ions (Kanavarioti *et al.* 2001; Monnard *et al.* 2003), as well as in hydrothermal environments (Ogasawara *et al.* 2000). Supposedly such spontaneous or catalytic activities may have been utilized for replication, genome processing and recombination of the early RNA replicons. Later, when protein-mediated catalysis replaced the early RNA functions, most of the RNA catalysts would have disappeared, while some apparently adopted new functions – such as the conversion of the ribosomes and tRNAs into the translation machinery. Some of the catalytic functions have also been maintained through the evolution, and occur still in DNA-based life forms, among the protein-mediated biochemistry. Some examples of such highly conserved RNA-catalysed functions are, at least, the small nucleolar RNA (snoRNA)-mediated processing of the ribosomal RNAs, the RNA-catalysed processing of the tRNAs by RNAse P in all three kingdoms of life, and autocatalytic, or spliceosome-mediated splicing of introns in eukaryotic genomes (Jeffares *et al.* 1998; Meli *et al.* 2001). Interestingly, these features are highly conserved and utilized in the eukaryotes, while several of them have been replaced by a corresponding protein-catalysed function in some prokaryotes (Maizels & Weiner 1994; Poole *et al.* 1998; Meli *et al.* 2001). In this respect the eukaryotic genomes seem to have retained the more original function, while the

https://doi.org/10.1017/S1473550405002521 Published online by Cambridge University Press

prokaryotes have adapted a more efficient catalysis for the same function. It can be assumed that evolution produced the more efficient, protein-mediated reaction pathways from the RNA-mediated slower ones, and not *vice versa*. Thus, the loss of RNA relicts from the prokaryotic metabolic pathways suggests that these genomes may have been derived from more ancestral, eukaryote-like genomes in the evolutionary process (Poole *et al*. 1998; Meli *et al*. 2001).

From replication to translation

In present-day life forms, translation is a very complex process involving many components (Fig. 1): mRNA provides the translatable nucleotide sequence, aminoacylated tRNAs bring the amino acids to the reaction and 23S rRNA (assisted by multiple ribosomal proteins) mediate the formation of the peptide bonds between amino acids (Zhang & Cech 1997; Steitz & Moore 2003; Steitz 2005). Each step of the translation process also requires many accessory proteins. Thus, translation in contemporary life presents a sophisticated, but precisely coordinated network of interacting RNA and protein components. Such a complicated translation apparatus and its functional mechanisms must have developed from a much simpler and primitive ancestor system, which could have existed in the RNA world, and this system had to rely on RNA–RNA interactions (Woese 2001). To allow the evolutionary development of complex functional structures, such as ribosomes and tRNAs, their primitive precursors needed to be replicated and selected (stepwise) for some pre-existing beneficial function. It has been proposed that the early function of the (core domains) of ribosome would have been the polymerization of nucleotides, maybe via ligation of nucleotide triplets, which were brought into the reaction by the anticodon stem of the (primitive) tRNA (Poole *et al*. 1998). Owing to the higher annealing energy, RNA replication via ligation of nucleotide triplets would have been much more efficient, and more accurate than addition of single nucleotides separately. Extension of the growing strand by nucleotide triplets could have also been the first step towards establishment of the three-nucleotide-based genetic code, and its interpretation (translation) via the recognition–binding reaction between triplet codons, and the corresponding anticodon loop of the aminoacyl tRNAs (Poole *et al*. 1998). The early role of ribosomes and of some contemporary translation factors in the RNA replication process is supported by the fact that still today, the ribosomal protein S1, and the translation elongation factors Tu and Ts are essential components of the replicase complex of Qβ bacteriophage (Maizels & Weiner 1994), and elongation factor(s) are also components of the replicase complexes of some plant RNA viruses (Ahlquist *et al*. 2003).

It is believed that the tRNA molecules were also crucial components of the initial replication system, as well as of the early machinery of protein synthesis (Brosius 2001; Woese 2001). According to the genomic tag hypothesis, tRNA-like molecules acted first as molecular tags to label those RNA molecules dedicated to serve as templates for RNA replication by RNA ribozymes (Weiner & Maizels 1987; Maizels & Weiner 1994). Also, tRNA aminoacylation could arise in such a system as an additional means to tag the replication substrates or as an intermediate in this tagging (Weiner & Maizels 1987). Not necessarily the whole tRNA structure was used as a tag for replication templates, but maybe only the 'top half' of modern tRNA, consisting of a coaxial stack of the TΨC arm on the acceptor stem, which is often considered the more ancient part of a tRNA, being the minimal substrate for chiral-selective aminoacylation (Maizels & Weiner 1994; Tamura & Schimmel 2004). Finally, aminoacyl tRNAs could become involved in protein synthesis via their interaction with the first protoribosomes, either as replication templates or as components of the replication machinery (Brosius 2001), as discussed above. Thus, it is possible that the early replication process *per se* led to the development of the translatable genetic code and the translation machinery (Jeffares *et al*. 1998; Poole *et al*. 1998).

The hypothesis of the early function of the tRNAs as a tag for replicating templates is based on the observation that tRNA-like moieties still occur in the telomerase enzyme complexes, and are used as a template of these enzymes in the re-synthesis, repair and maintenance of the ends of eukaryotic genomic DNAs (the telomeres). tRNAs are also used as primers of the DNA synthesis, when various retro- and pararetroviruses, or different retroelements convert their genomes from RNA to DNA form (Turner & Covey 1988; Maizels & Wiener 1994; McClure 1999; Weiner & Maizels 1999; Hull 2002). Many positive-strand RNA viruses, infecting bacteria and plants, contain a tRNA-like structure (TLS) at the 3′-end of the genomic RNA. TLSs of different viruses can be aminoacylated by histidine, valine or tyrosine (Dreher 1999). They function as primers for RNA replication, but in these systems, only the very 3′-terminal CCA or CCCA sequences of the TLS seem to be required for replication initiation (Yoshinari *et al*. 2000).

Owing to their apparent early emergence as the catalysts of the RNA world, and as the central components of the early replicons and of the early translation system, the present-day ribosomes and the tRNAs can be considered as molecular fossils of the RNA world, and as some of the most ancient relicts and catalytic units of the RNA-based life forms (Weiner & Maizels 1987, 1999; Poole *et al*. 1998). They have apparently played a central role in the development of the protein coding capacity (Stevenson 2002; Weberndorfer *et al*. 2003).

Models of early translation

Most mRNAs of contemporary eukaryotes possess the 5′-m^7G-cap structure and 3′-poly(A) tail. These elements are critically involved in translation initiation, which also depends on many protein translation initiation factors (eIFs), represented by at least 23 different polypeptides (Kapp & Lorsch 2004). However, some mRNAs of eukaryotes or their viruses are characterized by a cap-independent translation initiation (Merrick 2004). This may depend on ribosomal

https://doi.org/10.1017/S1473550405002521 Published online by Cambridge University Press

entry directly from the 5′-end of mRNA or, alternatively, be mediated by ribosomal loading onto the internal ribosome entry site (IRES), located in the 5′-non-translated region (NTR). IRES-mediated translation was originally discovered in picornaviruses (Jang *et al*. 1988; Pelletier & Sonenberg 1988) and flaviviruses (Tsukiyama-Kohara *et al*. 1992). IRES-dependent translation has different requirements for protein eIFs in different viruses (Hellen & Sarnow 2001; Martines-Salaz *et al*. 2001). Lower degree of protein dependence for IRES-mediated translation can be considered as evolutionary more primitive. For instance, translation initiation in picornaviruses, with the highly structured IRESs of more than 400 nt long, requires several of the same eIFs that are necessary for cap-dependent translation of cellular mRNAs, i.e. eIF4G, eIF4A, eIF3 and eIF2 (Bedard & Semler 2004; Martinez-Salas & Fernandez-Miragall 2004). A less protein-dependent mechanism is presented by flaviviruses, e.g. the hepatitis C virus (HCV), with extensively structured IRESs of about 350 nucleotides, which need only two eIFs (eIF2 and eIF3) for translation initiation (Sarnow 2003).

Even complete protein independence for IRES-mediated translation occurs in viruses from the *Dicistroviridae* family, where translation initiation is governed by two IRESs. Initiation on the 3′-IRES (intergenic region IRES, IGR-IRES), having a length of about 200 nt, occurs at the non-AUG codon, CUU for the *Plautia stali* intestinal virus (Sasaki & Nakashima 1999, 2000) or CCU for the cricket paralysis virus (CrPV; Wilson *et al*. 2000a). The IGR-IRES forms specific contacts with the ribosomal P (peptidyl) and E (exit) sites, inducing conformational changes in the ribosome, thus providing ribosomal binding, proper positioning of the viral mRNA inside the ribosome and triggering protein eIF-independent translation initiation and the first translocation without the preformed peptide bond (Wilson *et al*. 2000b; Jan & Sarnow 2002; Jan *et al*. 2003; Pestova & Hellen 2003; Cevallos & Sarnow 2005). In this way, by mimicking the functions of the P-site and E-site tRNAs and by active manipulating the ribosomal structural conformation, the IGR-IRES acts as a (relatively) protein-independent, RNA-based translation factor, covalently attached to the mRNA (Hatakeyama *et al*. 2004; Pestova *et al*. 2004; Spahn *et al*. 2004).

The TLSs located at the 3′ termini of viral RNAs can serve as translational enhancers, as shown for brome mosaic virus (Barends *et al*. 2004) and turnip yellow mosaic virus (Barends *et al*. 2003; Matsuda & Dreher 2004). In the case of the turnip yellow mosaic virus, the valylated viral TLS can direct ribosomes to the internal site in the 5′-part of the mRNA, forming the long-range RNA–RNA interaction between the TLS and the initiation site for translation of the viral polyprotein (Barends *et al*. 2003). This may lead to incorporation of valine as the first amino acid residue of the viral protein, independently of the initiation ternary complex and the cap-binding protein eIF4E. Therefore, the translation initiation mechanism used by the turnip yellow mosaic virus could be considered as reminiscent of the ancient relatively protein-independent protoribosome (Barends *et al*. 2003).

Curiously, the functionally important TLS domains have been found not only at the 3′-ends of viral mRNAs, but also in the IRES elements of different unrelated viruses, e.g. in members of the families *Picornaviridae*, *Flaviviridae* and *Dicistroviridae* (Jan *et al*. 2003; Lyons & Robertson 2003; Piron *et al*. 2005). Such tRNA-like domains are the only structural feature shared by these evolutionary distinct IRESs, but the exact functions of these domains remain to be determined.

Direct base-pairings between the 5′-NTR and different 3′-NTR sequences of viral mRNAs have been demonstrated to also be functionally important for translation initiation in the barley yellow dwarf virus (Guo *et al*. 2001), the tomato bushy stunt virus (Fabian & White 2004) and proposed for the tobacco necrosis virus (Meulewaeter *et al*. 2004; Shen & Miller 2004). Provided that these long-range RNA–RNA interactions function relatively independently of protein factors, they could be considered as functional relics of the primitive translation systems.

Although the phylogenetic relationships between eukaryotes and prokaryotes remain a subject of intensive debate (Jeffares *et al*. 1998; Poole *et al*. 1998; Meli *et al*. 2001), at least one prokaryotic mechanism is worth mentioning here. Transfer-messenger RNA (tmRNA) combines functional features of both tRNA and mRNA, being ubiquitous for bacteria and playing a central part in their metabolism. tmRNA, which can be aminoacylated with alanine, participates in a process called *trans*-translation, relieving the stalled ribosomes, increasing translational efficiency on rare codons and operating on certain stop codons (Muto *et al*. 1998; Withey & Friedman 2003). tmRNA is often considered as a functional model of an extant protoribosome (Brosius 2001; Meli *et al*. 2001; Di Giulio 2003).

Relicts of the RNA–protein world

The appearance of the first genetically encoded proteins and of the molecular machinery for protein synthesis gave the early replicons a huge advantage. Protein-mediated structural and catalytic functions have been so essential for the further development of life that the first hypothetical protein-encoding RNA organisms have been called the ‘breakthrough organisms’, named *Riborgis eigensis* (Jeffares *et al*. 1998; Poole *et al*. 1998; Meli *et al*. 2001).

It is likely that the first genetically encoded proteins were produced accidentally, through the interaction between the (polymerizing) ribosome advancing on a template RNA strand, and the aminoacyl tRNAs. The first translated polypeptides were most probably small, structurally simple and had no catalytic properties. However, they may have been able to interact with the RNA replicons, and may have improved their stability or functional conformation in a chaperone- or ligand-like manner, thus providing a selective function to drive their evolution (Poole *et al*. 1998; Noller 2004). The earliest catalytic functions, which evolved later, were most probably related to those functions, which were already established at that time, i.e. to the replication of the

https://doi.org/10.1017/S1473550405002521 Published online by Cambridge University Press

RNA genomes. Thus, it is possible that the RNA-dependent RNA polymerase may have been one of the original enzymatic proteins (Iyer *et al*. 2003). Later, DNA genomes may have arisen, possibly via the polymerization of the deoxynucleotides by the same enzyme (Maizels & Weiner 1994; Siegel *et al*. 1999) or via conversion of the enzyme activity towards reverse transcriptase. However, there is no statistically significant relationship between the reverse transcriptase and RNA-dependent RNA polymerase enzymes to support common ancestry (Zanotto *et al*. 1996; McClure 1999).

It can be assumed that the earliest enzyme-replicated RNA genomes, and their encoded protein products were very simple, and the first translatable templates coded only for a single polypeptide each. The primitive RNA polymerases were very error-prone and produced a high level of variation in the RNA sequences. These evolved to code for different polypeptides, potentially with new beneficial functions. All RNAs were still replicated by the same (or similar) replicase moieties, which recognized the same (or similar) replication tags (possibly tRNA-like) at the ends of the genomes. This type of replicon population would have formed a group of co-replicating, mutually beneficial linear genomes, tagged at the ends with tRNA-like (or telomere-like) structures. Such a replicon population would have been similar to the present-day structure of fragmented, linear eukaryotic genomes, further supporting the hypothesis that the eukaryotic genome structure may be more ancestral than the circular prokaryotic genomes.

One of the most significant restrictions for the proliferation, development and growth of the RNA-based genomes has been the inherent instability of the RNA polymers. RNA strands would hardly develop into large, stable and long-lived genomes, and therefore the development of DNA genomes had a strong selective advantage (Dworkin *et al*. 2003). However, RNA genomes could be significantly stabilized already by the invention of protein synthesis, because assembly of RNAs into proteinaceous complexes (RNP complexes) or encapsidation within protein structures can significantly increase their stability and extend their structural and, therefore, functional capabilities (Noller 2004). Also, RNAs can be stabilized by closure into membrane vesicles. Both of these strategies may have been important steps towards the development of protein-associated cell membranes.

Viruses and viroids as models of the RNP replicons

As discussed above, we can hypothesize that the early populations of RNP replicons were short linear RNAs, replicated either *in cis*, or *in trans*, by RNA-dependent RNA polymerase activities. The replicons may have contained various coding sequences which would have been mutually beneficial, some coding for the replicase, for example, and others coding for the protective coat protein or conformational subunits. This would have led to the development of interactive, fragmented genomes. Initially, the replicons would have been fully dependent on obtaining all of their building blocks (activated nucleotides and amino acids) from their immediate surrounding. They would have utilized very simple, RNA-regulated mechanisms for replication and translation. They would have evolved via error-prone RNA-dependent RNA polymerase, and by utilizing efficient RNA-recombination methods. Moreover, they would have improved their stability either by encapsidation into proteinaceous structures and/or by replicating in membraneous vesicles. Interestingly, all of these features are also shared by present-day RNA viruses. Thus, present-day viruses could be seen as 'functional relicts' of the early RNP replicons (Maizels & Weiner 1994).

Present-day viruses are molecular parasites: they can replicate only in environments (typically within their susceptible host cells) where they can obtain all of the molecules required for their replication. They are not primitive organisms: in suitable environments (inside of host cells) viruses replicate efficiently, and they can evolve and adapt to new environments (hosts). As they are fully heterotrophic, they are not viable on their own. Therefore, they are not alive, but rather on the border of being alive, in a similar way to the early replicons.

Many RNA viruses utilize maximally simple functions and molecular determinants for their replication, and for their translational regulation. Viruses also utilize deviant genetic codons to manipulate the translational machinery of the host cells (Hull 2002). Moreover, viruses utilize protein capsids, produced by a minimal amount of coding sequence for the protection of their RNAs. Several viruses replicate in specific vesicular structures formed by the host membranes (Lee & Ahlquist 2003), and some viruses (reoviruses) complete their replication process within the viral-encoded capsids (Fraenkel-Conrat & Kimball 1982; Grimes *et al*. 1998 and references therein), thus creating a cell-like separation from their environment. These features can be considered as putative models for the replication and survival strategies of the early replicons. Furthermore, some viruses (retro- and para-retroviruses) include in their life cycles both RNA- and DNA-based stages, by converting the RNA sequence by reverse transcription into DNA form, which may be reminiscent of the transition of the RNA genomes into DNA form (McClure 1999).

Viroids are also small RNA parasites that are fully dependent on their cellular hosts for replication and survival. They differ from RNA viruses in being smaller in size, varying from 246 to 401 nucleotides, circular, and not containing any protein coding sequences (Flores 2001). Thus, they do not contribute any proteins of their own for their replication, but instead, are replicated by the host-encoded RNA polymerases. They replicate via an intriguing rolling-circle mechanism, by producing multimeric copies of the viroid RNA, which are cleaved to unit-length RNAs by autocatalytic ribozyme sequences and then sealed to circular form either by RNA ligase or autocatalytically (Hutchins *et al*. 1986). In addition to their host-dependent replication and autocatalytic cleavage/ligation mechanisms, their survival strategy includes very intensive internal base-pairing of their genomic sequences. These internal secondary structures make

their genomes very stable against enzymatic or physical degradation, in spite of being naked RNAs, i.e. not associated with any protective proteins (Hull 2002; Flores *et al*. 2004).

It is conceivable that viroids may also resemble the early RNA replicons, in some respects even more so than the linear RNA viruses (Semancik & Duran-Vila 1999; Flores 2001). The ribozyme-mediated cleavage mechanism and base-pairing-mediated stability might have also been suitable survival strategies for the early replicons, and genome replication via the rolling-circle mechanism could have been an efficient way to achieve accurate replication of circular RNA genomes. The existence of these genomes without any coding sequences might also resemble the replicons of the era prior to the invention of protein synthesis. However, in this case the lack of coding capacity is combined with full dependence on host-encoded proteins. Thus, it does not seem to be related to self-maintaining RNAs, but rather, to highly evolved molecular parasitism. Of course, it is possible that molecular parasites were initiated along with those RNAs that functioned as replication and genome maintenance machineries. It seems that viroids might be compared to the former of these two groups.

From replicons to life

It is not clear what stage of the early evolution can be considered as the transition from molecular replication to life. Some authors suggest that life would have essentially started at the onset of molecular evolution (de Duve 1998; Joyce 2002), and this definition would identify the earliest replicons, driven by RNA catalysis, as the first life forms. Some authors consider that proteins, and their mediated structural and catalytic functions, are so essential in the development of any more advanced life forms that the first hypothetical protein-encoding RNA organisms have been called the 'breakthrough organisms' (Jeffares *et al*. 1998; Poole *et al*. 1998; Meli *et al*. 2001). However, according to present-day definitions replicating viral genomes, which code for their essential gene products but depend on their hosts (or their immediate environments) for all of their building blocks, are not considered alive (Hull 2002). According to this, the criteria for life would have been fulfilled only after a certain level of self-sustaining function had been reached. Cellular structure, molecular machineries and metabolism are essential for self-sustainable life. Therefore, one of the essential hallmarks of life may be containment in a cell membrane. The cell membrane separates the living entity from, but also mediates controlled communication and exchange of materials with the environment. Cell membranes bind the genomes and their beneficial gene products for joint evolutionary selection. Binary fission of membrane-contained cells facilitates concerted replication of fragmented genomes and membrane structures facilitate essential energy production pathways. The RNA replicons, and the RNA–protein-based breakthrough organisms would have been essential early developmental stages towards life, but life, as we know it, would have been established only after it was contained in cells, coordinated by DNA-encoded genetic information and catalysed by protein enzymes.

References

Ahlquist, P., Noueiry, A.O., Lee, W.-M., Kushner, D.B. & Dye, B.T. (2003). Host factors in positive-strand RNA virus genome replication. *J. Virol*. **77**, 8181–8186.

Barends, S., Bink, H.H.J., van den Worm, S.H.E., Pleij, C.W.A. & Kraal, B. (2003). Entrapping ribosomes for viral translation: tRNA mimicry as a molecular Trojan horse. *Cell* **112**, 123–129.

Barends, S., Rudinger-Thirion, J., Florentz, C., Giege, R., Pleij, C.W.A. & Kraal, B. (2004). tRNA-like structure regulates translation of *Brome mosaic virus* RNA. *J. Virol*. **78**, 4003–4010.

Bedard, K.M. & Semler, B.L. (2004). Regulation of picornavirus gene expression. *Microbes Infection* **6**, 702–713.

Brosius, J. (2001). tRNAs in the spotlight during protein biosynthesis. *Trends Biochem. Sci*. **26**, 653–656.

Cairn-Smiths, A.G. (1982). *Genetic Takeover and the Mineral Origins of Life*. Cambridge University Press, Cambridge.

Cevallos, R.C. & Sarnow, P. (2005). Factor-independent assembly of elongation-competent ribosomes by an internal ribosome entry site located in an RNA virus that infects penaeid shrimp. *J. Virol*. **79**, 677–683.

Chang, I.-F., Szick-Miranda, K., Pan, S. & Bailey-Serres, J. (2005). Proteomic characterization of evolutionary conserved and variable proteins of Arabidopsis cytosolic ribosomes. *Plant Physiol*. **137**, 848–862.

De Duve, C. (1998). Clues from present-day biology: the thioester world. In *The Molecular Origins of Life*, ed. Brack, A., pp. 219–236. Cambridge University Press, Cambridge.

Di Giulio, M. (2003). The early phases of genetic code origin: conjectures on the evolution of coded catalysis. *Origins Life Evol. Biosphere* **33**, 479–489.

Doudna, J.A. & Cech, T.R. (2002). The chemical repertoire of natural ribozymes. *Nature* **418**, 222–228.

Dreher, T.W. (1999). Functions of the 3′-untranslated regions of the positive strand RNA viral genomes. *Annu. Rev. Phytopathol*. **37**, 151–174.

Dworkin, J.P., Lazcano, A. & Miller, S.L. (2003). The roads to and from the RNA world. *J. Theor. Biol*. **222**, 127–134.

Fabian, M.R. & White, K.A. (2004). 5′–3′ RNA–RNA interaction facilitates cap- and poly(A) tail-independent translation of *Tomato bushy stunt virus* mRNA. *J. Biol. Chem*. **279**, 28 862–28 872.

Ferris, J.P. (2002). Montmorillonite catalysis of 30–50 mer oligonucleotides: laboratory demonstration of potential steps in the origin of the RNA world. *Origins Life Evol. Biosphere* **32**, 311–332.

Flores, R. (2001). A naked plant-specific RNA ten-fold smaller than the smallest known viral RNA: the viroid. *C.R. Acad. Sci. Paris, Sciences de la Vie/Life Sciences* **324**, 943–952.

Flores, R., Delgado, S., Gas, M.-E., Carbonell, A., Molina, D., Gago, S. & De la Pena, M. (2004). *FEBS Lett*. **567**, 42–48.

Fraenkel-Conrat, H. & Kimball, P.C. (1982). *Virology*. pp. 150–159. Prentice-Hall, Englewood Cliffs, NJ.

Gilbert, W. (1986). The RNA world. *Nature* **319**, 618.

Grimes, J.M., Burroughs, J.N., Gouet, P., Diprose, J.M., Malby, R., Zientara, S., Mertens, P.P. & Stuart, D.I. (1998). The atomic structure of the bluetongue virus core. *Nature* **395**, 470–478.

Guo, L., Allen, E.M. & Miller, W.A. (2001). Base-pairing between untranslated regions facilitates translation of uncapped, nonpolyadenylated viral RNA. *Mol. Cell* **7**, 1103–1109.

Hatakeyama, Y., Shibuya, N., Nishiyama, T. & Nakashima, N. (2004). Structural variant of the intergenic internal ribosome entry site elements in dicistroviruses and computational search for their counterparts. *RNA* **10**, 779–786.

Hellen, C.U.T. & Sarnow, P. (2001). Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev*. **15**, 1593–1612.

Hud, N.V. & Anet, F.A.L. (2000). Intercalation-mediated synthesis and replication: a new approach to the origin of life. *J. Theor. Biol*. **205**, 543–562.

Hull, R. (2002). *Matthew's Plant Virology*, pp. 339–345. Academic Press, London.

Hutchins, C.J., Rathjen, P.D., Forster, A.C. & Symons, R.H. (1986). Self-cleavage of plus and minus RNA transcripts of avocado sunblotch viroid. *Nucleic Acids Res*. **14**, 3627–3640.

Iyer, L.M., Koonin, E.V. & Aravind, L. (2003). Evolutionary connection between the catalytic subunits of DNA-dependent RNA polymerases and eukaryotic RNA-dependent RNA polymerases and the origin of RNA polymerases. *BMC Struct. Biol*. **3**, 1.

James, K.D. & Ellington, A.D. (1998). Catalysis in the RNA world. In *The Molecular Origins of Life*, ed. Brack, A., pp. 269–294. Cambridge University Press, Cambridge.

Jan, E. & Sarnow, P. (2002). Factorless ribosome assembly on the internal ribosome entry site of cricket paralysis virus. *J. Mol. Biol*. **324**, 889–902.

Jan, E., Kinzy, T.G. & Sarnow, P. (2003). Divergent tRNA-like element supports initiation, elongation and termination of protein biosynthesis. *Proc. Natl. Acad. Sci., USA* **100**, 15 410–15 415.

Jang, S.K., Krausslich, H.-G., Nicklin, M.J.H., Duke, G.M., Palmenberg, A.C. & Wimmer, E. (1988). A segment of the 5′ nontranslated region of *Encephalomyocarditis virus* RNA directs internal entry of ribosomes during *in vitro* translation. *J. Virol*. **62**, 2636–2643.

Jäschke, A. (2001). RNA-catalysed carbon–carbon bond formation. *Biol. Chem*. **382**, 1321–1325.

Jeffares, D.C., Poole, A.M. & Penny, D. (1998). Relicts from the RNA world. *J. Mol. Evol*. **46**, 18–36.

Johnston, W.K., Unrau, P.J., Lawrence, M.S., Glasner, M.E. & Bartel, D.P. (2001). RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extention. *Science* **292**, 1319–1325.

Joyce, G.F. (2002). The antiquity of RNA-based evolution. *Nature* **418**, 214–221.

Kanavarioti, A., Monnard, P.-A. & Deamer, D.W. (2001). Eutectic phases in ice facilitate nonenzymatic nucleic acid synthesis. *Astrobiology* **1**, 271–281.

Kapp, L.D. & Lorsch, J.R. (2004). The molecular mechanics of eukaryotic translation. *Annu. Rev. Biochem*. **73**, 657–704.

Kumar, R.K. & Yarus, M. (2001). RNA-catalyzed amino acid activation. *Biochemistry* **40**, 6998–7004.

Lathe, R. (2005). Tidal chain reaction and the origin of replicating biopolymers. *Int. J. Astrobiol*. **4**, 33–41.

Lee, W.-M. & Ahlquist, P. (2003). Membrane synthesis, specific lipid requirements, and localized lipid composition changes associated with a positive-strand RNA virus RNA replication protein. *J. Virol*. **77**, 12 819–12 828.

Lee, N., Bessho, Y., Wei, K., Szostak, J.W. & Suga, H. (2000). Ribozyme-catalysed tRNA aminoacylation. *Nature Struct. Biol*. **7**, 28–33.

Li, N. & Huang, F. (2005). Ribozyme-catalysed aminoacylation from CoA thioesters. *Biochemistry* **44**, 4582–4590.

Lyons, A.J. & Robertson, H.D. (2003). Detection of tRNA-like structure through RNaseP cleavage of viral internal ribosome entry site RNAs near the AUG start triplet. *J. Biol. Chem*. **278**, 26 844–26 850.

McClure, M.A. (1999). The retroid agents: disease, function and evolution. In *Origin and Evolution of Viruses*, ed. Domingo, E., Webster, R. & Holland, J., pp. 163–196. Academic Press, London.

Maizels, N. & Weiner, A.M. (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.

Martin, W. & Russell, M.J. (2002). On the origin of cells: a hypothesis for the evolutionary transformation from abiotic geochemistry to chemoauthotrophic prokaryotes, and from prokaryotes to nucleated cells. *Phil. Trans. R. Soc. Lond.* **B 358**, 59–85 (also published in www.gla.ac.uk/project/originoflife/html/2001/pdf_files/Martin_&_Russell.pdf).

Martinez-Salas, E. & Fernandez-Miragall, O. (2004). Picornavirus IRES: structure function relationship. *Curr. Pharm. Des*. **10**, 3757–3767.

Martinez-Salas, E., Ramos, R., Lafuente, E. & Lopes de Quinto, S. (2001). Functional interactions in internal translation initiation directed by viral and cellular IRES elements. *J. Gen. Virol*. **82**, 973–984.

Matsuda, D. & Dreher, T.W. (2004). The tRNA-like structure of Turnip yellow mosaic virus RNA is a 3′-translational enhancer. *Virology* **321**, 36–46.

Meli, M., Albert-Fournier, B. & Maurel, M.-C. (2001). Recent findings in the modern RNA world. *Int. Microbiol*. **4**, 5–11.

Merrick, W.C. (2004). Cap-dependent and cap-independent translation in eukaryotic systems. *Gene* **332**, 1–11.

Meulewaeter, F., van Lipzig, R., Gultyaev, A.P., Pleij, C.W.A., van Damme, D., Cornelissen, M. & van Eldik, G. (2004). Conservation of RNA structures enables TNV and BYDV 5′ and 3′ elements to cooperate synergistically in cap-independent translation. *Nucleic Acids Res*. **32**, 1721–1730.

Miller, S.L. (1998). The endogenous synthesis of organic compounds. In *The Molecular Origins of Life*, ed. Brack, A., pp. 59–85. Cambridge University Press, Cambridge.

Monnard, P.A., Kanavarioti, A. & Deamer, D.W. (2003). Eutectic phase polymerization of activated ribonucleotide mixtures yields quasiequimolar incorporation of purine and pyrimidine nucleobases. *J. Am. Chem. Soc*. **125**, 13 734–13 740.

Muto, A., Ushida, C. & Himeno, H. (1998). A bacterial RNA that functions as both a tRNA and an mRNA. *Trends Biochem. Sci*. **23**, 25–29.

Nakao, A., Yoshihama, M. & Kenmochi, N. (2004). RPG: the ribosomal protein gene database. *Nucleic Acids Res*. **32D**, 168–170.

Noller, H.F. (2004). The driving force for molecular evolution of translation. *RNA* **10**, 1833–1837.

Ogasawara, H., Yoshida, A., Imai, E.-I., Honda, H., Hatori, K. & Matsuno, K. (2000). Synthesizing oligomers from monomeric nucleotides in simulated hydrothermal environments. *Origins Life Evol. Biosphere* **30**, 519–526.

Orgel, L.E. (2004). Prebiotic chemistry and the origin of the RNA world. *Critical Rev. Biochem. Molec. Biol*. **39**, 99–123.

Pelletier, J. & Sonenberg, N. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **334**, 320–325.

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

Pestova, T.V. & Hellen, C.U.T. (2003). Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. *Genes Dev*. **17**, 181–186.

Pestova, T.V., Lomakin, I.B. & Hellen, C.U.T. (2004). Position of the CrPV IRES on the 40S subunit and factor dependence of IRES/80S ribosome assembly. *EMBO Rep*. **5**, 906–913.

Piron, M., Beguristain, N., Nadal, A., Martinez-Salas, E. & Gomez, J. (2005). Characterizing the function and structural organization of the 5′ tRNA-like motif within the hepatitis C virus quasispecies. *Nucleic Acids Res*. **33**, 1487–1502.

Poole, A.M., Jeffares, D. & Penny, D. (1998). The path from the RNA world. *J. Mol. Evol*. **46**, 1–17.

Ricardo, A., Carrigan, M.A., Olcott, A.N. & Benner, S.A. (2004). Borate minerals stabilize ribose. *Science* **303**, 196.

Robertson, H.G. & Neel, O.D. (1999). Virus origins: conjoined RNA genomes as precursors to DNA genomes. In *Origin and Evolution of Viruses*, ed. Domingo, E., Webster, R. & Holland, J., pp. 25–36. Academic Press, London.

Roossinck, M.J. (1997). Mechanisms of plant virus evolution. *Annu. Rev. Phytopathol*. **35**, 191–209.

Sarnow, P. (2003). Viral internal ribosome entry site elements: novel ribosome-RNA complexes and roles in viral pathogenesis. *J. Virol*. **77**, 2801–2806.

Sasaki, J. & Nakashima, N. (1999). Translation initiation at the CUU codon is mediated by the internal ribosome entry site of an insect picorna-like virus *in vitro*. *J. Virol*. **73**, 1219–1226.

Sasaki, J. & Nakashima, N. (2000). Methionine-independent initiation of translation in the capsid protein of an insect RNA virus. *Proc. Natl Acad. Sci., USA* **97**, 1512–1515.

Schmidt, J.G., Nielsen, P.E. & Orgel, L.E. (1997). Information transfer from peptide nucleic acids to RNA by template-directed syntheses. *Nucleic Acids Res*. **25**, 4797–4802.

Schwartz, A.W. (1998). Origin of the RNA world. In *The Molecular Origins of Life*, ed. Brack, A., pp. 237–254. Cambridge University Press, Cambridge.

Semancik, J.S. & Duran-Vila, N. (1999). Viroids in plants: shadows and footprints of a primitive RNA. In *Origin and Evolution of Viruses*, ed. Domingo, E., Webster, R. & Holland, J., pp. 37–64. Academic Press, London.

Shen, R. & Miller, W.A. (2004). The 3′ untranslated region of *Tobacco necrosis virus* RNA contains a *Barley yellow dwarf virus*-like cap-independent translation element. *J. Virol*. **78**, 4655–4664.

Siegel, R.W., Bellon, L., Beigelman, L. & Kao, C.C. (1999). Use of DNA, RNA and chimeric templates by a viral RNA-dependent RNA polymerase: evolutionary implications for the transition from the RNA to DNA world. *J. Virol*. **73**, 6424–6429.

Spahn, C.M.T., Jan, E., Mulder, A., Grassucci, R.A., Sarnow, P. & Frank, J. (2004). Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. *Cell* **118**, 465–475.

Steitz, T.A. (2005). On the structural basis of peptide-bond formation and antibiotic resistance from atomic structures of the large ribosomal subunit. *FEBS Lett*. **579**, 955–958.

Steitz, T.A. & Moore, P.B. (2003). RNA, the first macromolecular catalyst: the ribosome is a ribozyme. *Trends Biochem. Sci*. **28**, 411–418.

Stevenson, D.S. (2002). Co-evolution of the genetic code ans ribozyme replication. *J. Theor. Biol*. **217**, 235–253.

Tamura, K. & Schimmel, P. (2004). Chiral-selective aminoacylation of an RNA minihelix. *Science* **305**, 1253.

Tsukiyama-Kohara, K., Luzika, N., Kohara, M. & Nomoto, A. (1992). Internal ribosome entry site within hepatitis C virus RNA. *J. Virol*. **66**, 1476–1483.

Turner, D.S. & Covey, S.N. (1988). Discontinuous hairpin DNAs synthesized *in vivo* following specific and non-specific proming of cauliflower mosaic virus DNA (+) strands. *Virus Res*. **9**, 49–62.

Veuthey, A.L. & Bittar, G. (1998). Phylogenetic relationships of fungi, plantae, and animalia inferred from homologous comparison of ribosomal proteins. *J. Mol. Evol*. **47**, 81–92.

Vlassov, A.V., Johnston, B.H., Landweber, L.F. & Kazakov, S.A. (2004). Ligation activity of fragmented ribozymes in frozen solution: implications for the RNA world. *Nucleic Acids Res*. **32**, 2966–2974.

Wächtershäuser, G. (1998). Origin of life in an iron–sulfur world. In *The Molecular Origins of Life*, ed. Brack, A., pp. 206–218. Cambridge University Press, Cambridge.

Weberndorfer, G., Hofacker, I.L. & Stadler, P.F. (2003). On the evolution of primitive genetic codes. *Origins Life Evol. Biosphere* **33**, 491–514.

Weiner, A.M. & Maizels, N. (1987). tRNA-like structures tag the 3′ ends of genomic RNA molecules for replication: implications for the origin of protein synthesis. *Proc. Natl Acad. Sci., USA* **84**, 7383–7387.

Weiner, A.M. & Maizels, N. (1999). The genomic tag hypothesis: modern viruses as molecular fossils of ancient strategies for genomic replication, and clues regarding the origin of protein synthesis. *Biol. Bull*. **196**, 327–330.

Wilson, J.E., Powell, M.J., Hoover, S.E. & Sarnow, P. (2000a). Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites. *Mol. Cell. Biol*. **20**, 4990–4999.

Wilson, J.E., Pestova, T.V., Hellen, C.U.T. & Sarnow, P. (2000b). Initiation of protein synthesis from the A site of the ribosome. *Cell* **102**, 511–520.

Withey, J.H. & Friedman, D.I. (2003). A salvage pathway for protein synthesis: tmRNA and trans-translation. *Annu. Rev. Microbiol*. **57**, 101–123.

Woese, C.R. (1998). The universal ancestor. *Proc. Natl Acad. Sci., USA* **95**, 6854–6859.

Woese, C.R. (2001). Translation: in retrospect and prospect. *RNA* **7**, 1055–1067.

Yoshinari, S., Nagy, P.D., Simon, A.E. & Dreher, T.W. (2000). CCA initiation boxes without unique promoter elements support *in vitro* transcription by three viral RNA-dependent RNA polymerases. *RNA* **6**, 698–707.

Zanotto, P., Gibbs, M.J., Gould, E.A. & Holmes, E.C. (1996). A reevaluation of the higher taxonomy of viruses based on RNA polymerases. *J. Virol*. **70**, 6083–6093.

Zhang, B. & Cech, T.R. (1997). Peptide bond formation by *in vitro* selected ribozymes. *Nature* **390**, 96–100.

Zubay, G. & Mui, T. (2001). Prebiotic synthesis of nucleotides. *Orig. Life Evol. Biosph*. **31**, 87–102.