Hypotheses

The RNA dreamtime

Modern cells feature proteins that might have supported a prebiotic polypeptide world but nothing indicates that RNA world ever was

Charles G. Kurland

Modern cells present no signs of a putative prebiotic RNA world. However, RNA coding is not a *sine qua non* for the accumulation of catalytic polypeptides. Thus, cellular proteins spontaneously fold into active structures that are resistant to proteolysis. The law of mass action suggests that binding domains are stabilized by specific interactions with their substrates. Random polypeptide synthesis in a prebiotic world has the potential to initially produce only a very small fraction of polypeptides that can fold spontaneously into catalytic domains. However, that fraction can be enriched by proteolytic activities that destroy the unfolded polypeptides and regenerate amino acids that can be recycled into polypeptides. In this open system scenario the stable domains that accumulate and the chemical environment in which they are accumulated are linked through self coding of polypeptide structure. Such open polypeptide systems may have been the precursors to the cellular ribonucleoprotein (RNP) world that evolved subsequently.

Keywords:
domain selection; non-ribosomal peptidyl transferase; polypeptides; proteolysis; ribozymes

Introduction

The standard model for the origin of life imagines that the first replication, translation, and transcription systems were supported by RNA without the intervention of proteins [1]. The impulses for this conjecture were first, the discoveries of introns and exons in eukaryote mRNAs and second, the self splicing of some intron sequences in ribosomal RNA [2]. Here, a core assertion is that primordial mini-RNAs corresponding to the original exons encoded the first polypeptides [1]. In this scenario ''RNA molecules began to synthesize proteins, first by developing RNA adapter molecules that can bind activated amino acids and then by arranging them according to an RNA template using other molecules such as the RNA core of the ribosome''. Twenty-four years later this dazzling speculation has been reduced by ritual abuse to something like a creationist mantra. Hence, the title, borrowed from Collins *et al*. [3], alludes to an oral tradition of origins passed on by the first Australians. Finally, the support for a prebiotic RNA world consists solely of ingenious piecemeal chemical simulations *in vitro* that were obtained by chemists at great cost and effort over a 20-year period [4]. Such chemical simulations are accepted as hard evidence by RNA worlders, but in truth, they do not constitute proper evolutionary evidence.

It might have been useful earlier on to address questions such as: Does the routine identification of ribonucleoproteins (RNPs) as remnants of a prebiotic RNA world [4] support Gilbert's conjecture [1] or do they beg the question? Is partial simulation *in vitro* by a ribozyme running at one-millionth the rate normally catalyzed by a protein [5, 6] to be taken as evidence for the prebiotic precedence of the ribozyme activity? Though there are impressive synthetic ribozymes with convincing performance characteristics [7], why are there no examples of naturally occurring, protein-free ribozymes to link the postulated protein-free RNA world to the modern cellular world [8, 3]?

The secret of the cage

RNase P is an RNP that mediates the maturation of transfer RNAs [9]. Normally, RNase P has both RNA as well as protein components. But recently, an RNA-free variant was discovered in human mitochondria and shown to

DOI 10.1002/bies.201000058

Department of Microbial Ecology, University of Lund, Sölvegatan, Lund, Sweden

Corresponding author:
Charles G. Kurland
E-mail: cgkurland@yahoo.se

Abbreviations:
FSF, fold superfamily; **NPTase**, non-ribosomal peptidyl transferase; **PTC**, peptidyl transferase center; **RNP**, ribonucleoprotein.

mature tRNA precursors normally [10]. If RNase P function can be supported by protein alone, it is conceivable that such protein functions have participated in prebiotic systems as well. So, which came first, a protein or an RNP version of the enzyme? In fact, a recent phylogenetic survey suggests that an ancestral protein version may have preceded the RNP version of RNase P [11].

Another instructive RNP is the bacterial ribosome. Studies of these RNPs reveal the tension between routine biochemical observations and the expectations of RNA worlders, who refuse to accept the simplest interpretations of their experiments. For example, gentle disruption of 50S subunits yields a compacted 23S RNA associated with a very small core of ribosomal protein and this RNP supports peptidyl transferase activity with model substrates [12, 13]. Though these are straightforward results, the authors are reluctant to concede the potential involvement of ribosomal protein in the peptidyl transferase activity. Instead, they speculate that the core of proteins in their active 23S RNA preparations is trapped in an imaginary RNA cage; that is to say, the proteins are passive prisoners and not active participants [12, 13]. Noller [13] further comments that ''true 'ribocentrics' will simply view the latest aspect of the ribosomal puzzle as a worthy challenge, whose solution promises to reveal one of nature's most ancient biological secrets''.

An attempt to simulate the peptidyl transferase center (PTC) with the aid of a synthetic ribozyme was equally obtuse. Variants of tailored polynucleotides that could simulate peptidyl transferase activity were synthesized and selected *in vitro* [6]. The fastest variant of the tailored ribozymes had a k_{cat} corresponding to 0.05 minute^{-1} per peptide, which was compared with an elongation rate of 15–20 seconds^{-1} for *Escherichia coli* ribosomes [6]. In this comparison the k_{cat} was at least four orders of magnitude smaller for the ribozyme than for the ribosome. However, the k_{cat} of *E. coli* ribosomes in the peptidyl transferase reaction is closer to 100 seconds^{-1} [14], which indicates that the ribozyme is five orders of magnitude slower than the bacterial ribosome. Nevertheless, Zhang and Cech [6] suggest that the ''peptidyl transferase reaction of the selected ribozymes is 'fundamentally' similar to that carried out by the ribosome''. Fundamentally similar?

The first crystallographic reports of the structure of archaeal 50S ribosomal subunits described a roughly 20 Å-diameter protein-free RNA domain that was identified as the PTC [15, 16]. These observations contrasted decades of biochemistry that had identified protein contributions to the bacterial PTC [17].

So, a *fatwa* was issued to clarify the X-ray revelations: ''From this structure they deduced ... that RNA components of the large subunit accomplish the key peptidyl transferase reaction ... Thus, ribosomal RNA (rRNA) does not exist as a framework to organize catalytic proteins. Instead, the proteins are the structural units and they help to organize key ribozyme (catalytic RNA) ... elements, an idea long championed by Harry Noller, Carl Woese ..., and others'' [18].

In unambiguous contrast to this claim, Zimmerman and collaborators [17] showed that deletion of *E. coli* ribosomal protein L27, long thought to be part of the bacterial PTC, severely depresses bacterial growth rates, which are restored when the protein is expressed from a plasmid. Likewise, deletion of one to three N-terminal amino acids of L27 had a similar debilitating effect on growth rates as well as on peptidyl transferase reaction rates *in vitro*. Most informative was their finding that deletion of the three N-terminal amino acids strongly inhibits the labeling of L27 by photo-activated tRNA at the P site. The experiments of Maguire *et al.* [17] rather clearly confirm a substantial chain of biochemical experiments initiated in 1973 that consistently implicated L27 as a close neighbor and probable participant in the peptidyl transferase reaction of bacterial ribosomes.

The identification by Maguire *et al.* [17] of L27 as an essential part of the PTC in *E. coli* ribosomes was confirmed by a high-resolution structure of 70S active ribosomes with tRNA-filled A site and P site that revealed two proteins interacting directly with tRNA at the putative PTC [19]. One of these is L27 and the other is L16, which was also previously implicated in peptidyl transferase activity [17]. Thus, the universal protein-free ribozyme at the heart of the ribosome is history.

Does RNA replace protein?

Comparisons with the findings from archaeal subunits [15, 16] are suggestively complicated by the apparent absence of an L27 homolog from the archaeal 50S ribosomal subunit. According to Voorhees *et al.* [19] the projected orientation of L16 in the archaeal 50S subunits suggests that, like its homolog in the *E. coli* ribosome, it may be interacting with the elbow of the tRNA in the archaeal A site. In fact it would be highly instructive if, after more stringent structural determinations are made, archaeal ribosomes were indeed found to exploit a protein-free RNA peptidyl transferase in a high-resolution structure for active 70S particles. In that case, the PTC would be just another example of an RNP featuring an interchangeability of RNA and protein functions as in the RNA-free RNase P [10]. Indeed a cryptic evolutionary trade-off between RNA and protein may account for the observed differences between some archaeal and bacterial PTCs.

The growth efficiency of a cellular process such as translation or transcription can be measured by its rate normalized to the molecular mass that is required to carry out that process [20]. No biosynthetic cycle is as expensive to a prokaryote cell as translation because nothing else involves as large an investment in macromolecular equipment. This means that the biosynthetic cost of RNA in terms of growth efficiency is roughly one tenth that for a polypeptide chain. In this context of growth efficiency, the observed differences between some bacterial and archaeal 50S subunits are consistent with the notion that the evolution of the archaeal translation apparatus involved a trade-off between the costs of making that equipment and the efficiency with which it works.

For cells such as the bacterium *E. coli*, the growth optimization favors high rates of function in relatively rich media. That is to say, a fast, protein-rich PTC is an acceptable ribosomal design for optimal growth under relatively generous conditions [20, 21]. However, for cells such as some archaea the options may be different because they are adapted to growth under conditions of energy stress [22]. Here, ''cheap'' structural solutions for ribosomes are at a

premium even if they might come at the cost of a somewhat slower rate of function, as expected for ribozymes. The point is that archaea growing under conditions of energy stress are constrained to produce amino acids relatively slowly with the consequence that optimal translation rates might be correspondingly slow. Here, substitution of costly protein by less costly RNA would be a favored design strategy. Indeed, comparative data suggest that the evolution of the archaea involves a selective loss of proteins from ribosomes, a loss that is most striking at the crown of the archaeal ribosome tree [23].

Valentine's [22] insight into the adaptive specializations of the archaea may account for the putative protein-free PTC of some archaeal ribosomes, and more generally, it may explain why there are two prokaryote domains that are descendents of the eukaryote ancestor: one for rich environments and one for more metabolically challenging circumstances. Since the eukaryotes are the ancestral lineage from which the divergence of archaea and bacteria are thought to have been driven by reductive pressure [24–27], relaxed reductive pressure would allow eukaryote ribosomes to remain more proteinaceous than prokaryote ribosomes.

Accordingly, one prediction of this scenario is that the PTC of eukaryote ribosomes will turn out to be more protein-rich than those of prokaryotes. Another is that the rates of protein-depleted archaeal ribosomes under comparable conditions may prove to be slower than those of bacterial ribosomes. Astonishingly, it is currently impossible to locate in the literature ribosomal rates of translation along with growth rates with which to make meaningful comparisons.

Ribosomes are ribonucleoprotein particles, period!

An evolutionary trade-off between RNA and protein for the adaptations of ribosomes to the realities of growth constraints is consistent with an earlier view of ribosome structure. Voorhees *et al.* [19] note that even in *E. coli* ribosomes the PTC seems to be rich in RNA. This was not too surprising since "everywhere" in the bacterial ribosome is rich in RNA. Prior to the revelation of RNA world, an emergent idea was that ribosomal RNA and proteins are not segregated but are intermixed with cooperative clusters of proteins organized around specific RNA domains [28]. This was a data-driven interpretation based on biochemical studies of ribosomes modified by mutations, antibiotics, cross-linking agents, site-specific chemical agents, and partial assembly *in vitro*. The *E. coli* ribosome along with all other ribosomes was seen as a RNP particle, not as an RNA particle in protein drag.

Here, the self assembly of ribosomes generates mixed neighborhoods of RNA and proteins that provide the working surfaces for the translating ribosome [28]. The fact that a protein and an RNA domain cooperate in the assembly process would not preclude either one as a potential ligand for intermediates in translation. In effect, the strict division of labor for ribosomal RNA and protein proclaimed by Cech [18] is not necessarily respected by ribosomes. Amino acyl-tRNAs, protein factors, and mRNAs are all relatively huge substrates. Accordingly, the binding of these substrates to ribosomal sites and their movements during translation span correspondingly large RNP domains [28]. This data-driven, low-resolution model has been well substantiated by high-resolution structural data for the components of bacterial ribosomes. For example, all proteins with the exception of the oligomeric protein L7/L12 are directly bound to ribosomal RNA [29]. Likewise, participation of RNA and protein at functional sites is confirmed by high-resolution structures for 70S ribosomes complexed with A-site and P-site tRNAs [19].

Polypeptide world

Gilbert's "big bang" scenario [1] can be replaced by a data-driven, gradualist scheme in which a prebiotic polypeptide world evolved into a modern RNP world. Intelligent discussions of a putative polypeptide world are found in Cairns-Smith [30], Kaufman [31], and Egel [32]. Here, the principle novelty is a chemical scheme in which randomly generated, catalytic polypeptides may have been selected through a proteolytic mechanism that enriches the population of polypeptides with biologically relevant activities without the intervention of coding by RNA.

Highly relevant to this enrichment scheme are the workings of a ubiquitous cellular catabolic pathway that preserves the stability of high-density cellular proteomes by destroying proteins that present aberrant sequences [33–35]. Modern proteins are made up of one to several compact or folded domains (*e.g.* fold superfamilies or FSFs) along with terminal as well as interspersed linker sequences. The provision of robust folding pathways for polypeptides synthesized on ribosomes is part of normal sequence selection in evolution [36]. The result is that linkers are bound at surfaces of domains or to other macromolecules so that they, together with the self-organizing domains, are protected from the depredations of ubiquitous proteolytic "machines" such as proteasomes [33–35]. In general, degrading enzymes require a sequence that is less than 10 Å in diameter to be accommodated within the proteolytic site [37]. Clearly, an amino acid sequence that is organized into a compact domain or one that is stably bound to ligands such as other domains, lipid membranes, or chemical substrates is not likely to pass through the 10 Å gateway leading to proteolysis.

The presentation of proteolytically accessible amino acid sequences either as unbound linkers or as unfolded domains can result from mutation, translation errors, and chemical modification, induced conformational rearrangements, or failure to bind a ligand [33–36]. Systematic destruction of aberrant proteins by proteasomes and their homologs rids cellular proteomes of potential seeds for aggregation and precipitation, which otherwise would be lethal to cells. Indeed, the maintenance of the high protein densities characteristic of all cellular proteomes has had a profound influence on the evolution of cells [38, 26]. Human degenerative diseases that arise from mutations affecting protein folding or from defective proteasome function underscore the impact of proteolytic surveillance systems [39–42]. The metabolic consequence of such proteolytic surveillance is that protein degradation is a significant catabolic

flow in healthy cells. As much as thirty percent of newly synthesized proteins are destroyed by proteolysis [43], while a steady-state background of protein turnover of about two percent per hour is observed in growing cells [44].

This ubiquitous catabolic flow was introduced to illustrate a particularly informative dimension of protein evolution: this is that mutant variants presenting unfolded domains or unprotected strings would tend to be culled from populations as lethal alleles [45]. For example, even an allele that is expressed with a fully active catalytic site in an unstable fold might be a lethal allele if its substrate does not protect the domain from destruction. Accordingly, protein sequence evolution is not a random walk through amino acid sequence space. It is canalized through modular sequences that are expressed as self-organized compacted domains and protected linker sequences, which are the sole survivors of selection by proteolytic machines [45]. Indeed, such modular domains may correspond to the products of exons though it must be said that protein chemists insist that the boundaries of exons and domains are not identical [46]. Perhaps a measure of sequence drift over time may account for these boundary discrepancies.

Numerous enzymes that synthesize oligopeptides without the assistance of an mRNA or the rest of the modern translation apparatus have been described in all three superkingdoms [47–49]. Their products are short peptides with activities as diverse as antibiotics as well as neural transmitters and that ubiquity itself speaks for their ancient origins. Though these non-ribosomal peptidyl transferases (NPTases) produce oligopeptides with defined amino acid sequences, more primitive ancestral enzymatic activities lacking such amino acid specificity can be expected to have arisen in stochastic populations of polypeptides assembled by geochemical mechanisms [30–32]. So, the second tier of prebiotic polypeptides may have been composed of random sequences produced by NPTases. In this second phase, NPTases would be able to autocatalytically increase the rates with which random polypeptide sequences could accumulate.

Among sufficiently large populations of random amino acid sequences, some small fraction would be expected to fold spontaneously into stable active domains with functions such as the NPTases and proteolytic enzymes as well as any number of ''metabolic'' activities. This expectation is supported by the observations of Keefe and Szostak [50], who recovered biological activities from relatively small populations of random amino acid sequences polymerized *in vitro*. Thus, stochastic populations of polypeptides containing NPTases and proteolytic activities *etc*. could generate a dynamic situation in which random polypeptides are continuously synthesized, but most of these would be recycled by proteases that regenerate the amino acids. The small core of polypeptides that is resistant to proteolysis would be, by hypothesis, that which could spontaneously fold into stable compacted domains, which are resistant to enzymatic attack. These diverse domains, stable to proteolysis, would be enriched for a corresponding diversity of catalytic activities. This proposal can be tested in systems such as those described by Keefe and Szostak [50] to determine whether exposure of randomly synthesized polypeptides to proteolysis increases the specific activities of the polypeptides for assayable functions.

Finally, it is reasonable to expect that folding into active, proteolysis-resistant domains is facilitated by the binding of cofactors and/or substrates specific to individual catalytic polypeptides. Here, the specific binding of small molecules from the geochemical system to their cognate polypeptides would tend through the law of mass action to trap the polypeptides in a folded state that is resistant to proteolysis. In this way, prebiotic geochemistry may have selected polypeptides that mediate cycles of metabolic intermediates. The predicted effect of substrates on the proteolytic stabilities of random polypeptides can also be studied *in vitro* as above.

Three steps to cells

In fact, a prebiotic polypeptide scenario might account for the origins of the ribonucleotides that accumulated prior to the debut of RNA. There has been a tortured history of attempts to generate ribonucleotides in the laboratory as precursors required for the RNAs of a putative prebiotic world [51, 52]. Recently, an ingenious approach has succeeded to synthesize pyrimidine ribonucleotides in a reaction that might conceivably have been supported by volcanic outgassing [52]. However, there are still questions about the feasibility of this reaction scheme in a prebiotic environment and it still remains to account for the purine ribonucleotides [51, 52]. Finally, the scheme of Powner *et al*. [52] is not the textbook pathway for cellular biosynthesis of ribonucleotides from three amino acids. For these reasons it is still worth considering the alternative in which the ribonucleotides were accumulated prebiotically according to the textbook schemes that modern polypeptides follow.

Several innovations are required to enable a transition from the prebiotic world of polypeptides to biological systems. One is the introduction of an mRNA analog to encode proteins, and associated with that an analog of the aminoacyl-tRNA adapter to translate mRNAs, as in Gilbert's proposal [1]. However, the adapter initially need not have been a tRNA because the ability to recognize and to bind nucleotide triplets in mRNA is not unique to RNA. Proteins in the form of release factors can bind and recognize triplet codons with a discriminatory capacity much greater than that of aminoacyl-tRNAs [53]. Likewise, proteins in the form of aminoacyl-tRNA synthases recognize subtle side chain differences of amino acids and match these with cognate tRNA structures. Without this polypeptide function the adapter hypothesis would be just another RNA world fantasy.

So, whatever the disadvantages may be, a primitive translation machine may have exploited proteins for all functions except that of mRNA. That is to say, the evolution and tuning of the translation machinery may have been driven by the introduction and progressive expansion of RNA functions at the expense of protein. Such a tendency is consistent with the notion that the large protein-rich eukaryote ribosome is ancestral to the more economic archaeal ribosomes that have evolved with reduced protein complements under stringent reductive pressures ([24–27], Wang *et al*., in preparation).

A second elementary innovation would entail the ability to copy mRNA from complementary polynucleotides. A conservative guess is that this innovation arose when polypeptide-dependent random synthesis of polyribonucleotides was transformed into an enzymatic copy mechanism. Initially, it is simplest to imagine that making mRNA copies as transcripts and replicating them as genomes were one and the same function. At some point these two functions were separated as in modern cells.

The third innovation, which is essential for the evolution of genetically determined sequences, is the creation of a link between individual genetic determinants and their products to enable selection of competitive characters. Evolution has chosen cellular boundaries to provide this link. As a result, the composite features of a cell's proteome can be selected by their integrated influence on, for example, a cell's growth phenotype. Without such a link, selection is impossible. Obviously, selection was for this very reason impossible in the original formulation of RNA world [1].

There are at least two ways to think about the origins of cellular boundaries. One is to imagine that short polypeptides with amphiphilic character that could mimic lipid molecules formed the first membranous boundaries as for example in nanotubules made in laboratories [32, 54, 55]. The other is to employ lipids synthesized by polypeptides to spontaneously form vesicles. In appropriate solutions lipids associate and behave astonishingly like cellular membranes in model experiments [56–58]. Of course some combination of both lipids and amphiphilic peptides may be more relevant. Since the lipids as well as lipid-like peptides are potentially products of polypeptide enzymes, their emergence in the prebiotic world would not be exceptional.

Since each of these three innovations is in the present view a spontaneous expression of protein chemistry, it is assumed that the order of their appearances was random. However, all three innovations would have to come together to enable the transition from prebiotic to biological systems.

Conclusions

In the 1970s I attended an EMBO Workshop at which a French philosopher of science asserted that Molecular Biology was merely "romantic idealism". I of course indignantly rejected that comment out of hand. But 30 years later while trying to understand the explanatory power of Gilbert's "big bang" theory [1] for the origin of life I drifted back to that comment about romantic idealism. And, I began to understand why RNA world did not need to explain anything in order to be attractive to nearly all molecular biologists.

RNA world is an expression of the infatuation of molecular biologists with base pairing in nucleic acids played out in a one-dimensional space with no reference to time or energy: "DNA makes RNA makes protein" [59]. This is not chemistry. It is genetics. And, when true believers apply their genetic dogma to studies of chemical mechanism, the result is "the secret of the cage" and a five order of magnitude kinetic discrepancy described as a "fundamental" similarity [12, 13, 6].

The positive side of this infatuation has been the development of robust genomics, especially bioinformatics. But RNA worlders are not likely to find much comfort in genome sequences. A very recent phylogenomic study by Caetano-Anolles *et al.* [60] based on hundreds of fully sequenced genomes has revealed a time line for cellular evolution in which protein domains (FSFs) that support conventional metabolic pathways precede the debut of the nucleic acids as well as the protein domains associated with nucleic acids in gene expression. This cellular time line for the gradual elaboration of RNA functions [60] is not inconsistent with the thesis that the prebiotic world was a polypeptide world.

Protein folding into compact domains is a kind of self coding. However, the protein code is not a simple iterative code that lends itself to a copy mechanism. That is to say, it would be difficult to build genetics around polypeptide interactions alone. On the other hand, self folding and proteolytic editing might be just good enough to create a prebiotic chemical platform from which a cellular genetic system might take off.

If my own initial reactions are anything to go by, RNA world or "RNA makes RNA makes protein" has immediacy for molecular geneticists that is lacking in the present scheme for a prebiotic polypeptide world. Thus, the appeal of a polypeptide fold editing scheme rests on some familiarity with the contours of protein structure as well as with translational editing in modern cells. Regardless of its spontaneous appeal, I suggest that RNA world should now take its place on the shelf of "nice ideas" along with Aristotle's identifications of whales as fish and the worker bee as a male.

Acknowledgments

For stimulation, critique, unpublished information, and guidance in the literature I am greatly indebted to O. Berg, G. Caetano-Anolles, L. Collins, M. Ehrenberg, R. Garrett, A. Harish, A. Liljas, D. Penny, D. Valentine, and M. Wang. My gratitude also goes to Irmgard Winkler for help with the manuscript and to the Royal Physiographic Society, Lund, as well as the Nobel Committee for Chemistry at the Royal Swedish Academy of Sciences, Stockholm, for generous support.

References

1. **Gilbert W.** 1986. The RNA world. *Nature* **319**: 618.
2. **Zaug AJ**, **Cech TR.** 1986. The intervening sequence RNA of Tetrahymena is an enzyme. *Science* **231**: 470–5.
3. **Collins LJ**, **Kurland CG**, **Biggs P**, **Penny D**. 2009. The modern RNP world of eukaryotes. *J Hered* **100**: 597–604.
4. **Gesteland RF**, **Cech TR**, **Atkins JF.** ed. 2006. *The RNA World*. Woodbury, NY: Cold Spring Harbor Laboratory Press.
5. **Jeffares DC**, **Poole AM**, **Penny D.** 1998. Relics from the RNA World. *J Mol Evol* **46**: 18–36.
6. **Zhang B**, **Cech TR.** 1997. Peptide formation by *in vitro* selected ribozymes. *Nature* **390**: 96–100.
7. **Shechner DM**, **Grant RA**, **Bagby SC**, **Koldobskaya Y**, *et al*. 2009. Crystal structure of the catalytic core of an RNA-polymerase ribozyme. *Science* **326**: 1271–5.
8. **Collins L**, **Penny D.** 2009. The RNA infrastructure: dark matter of the eukaryotic cell? *Trends Genet* **25**: 120–8.
9. **Guerrier-Takada C**, **Lumelsky N**, **Altman S.** 1989. Specific interactions in RNA enzyme-substrate complexes. *Science* **246**: 1578–84.

10. **Holzmann J**, **Frank P**, **Löffler E**, **Bennett KL**, *et al*. 2008. RNaseP without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell* **135**: 462–4.
11. **Sun FJ**, **Caetano-Anolles G.** 2010. The ancient history of the structure of ribonuclease P and the early origins of Archaea. *BMC Bioinformatics* **11**: 153.
12. **Noller HF**, **Hoffarth V**, **Zimniak L.** 1992. Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* **256**: 1416–9.
13. **Noller HF.** 1993. Peptidyl transferase: protein, ribonucleoprotein, or RNA? *J Bacteriol* **175**: 5297–300.
14. **Johansson M**, **Lovmar M**, **Ehrenberg M.** 2008. Rate and accuracy of bacterial protein synthesis revisited. *Curr Opin Microbiol* **11**: 1–7.
15. **Ban N**, **Nissen P**, **Hansen J**, **Moore PB**, *et al*. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**: 905–20.
16. **Nissen P**, **Hansen J**, **Ban N**, **Moore PB**, *et al*. 2000. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**: 920–30.
17. **Maguire BA**, **Beniaminov AD**, **Ramu H**, **Mankin AS**, *et al*. 2005. A protein component at the heart of an RNA machine: the importance of protein L27 for the function of the bacterial ribosome. *Mol Cell* **20**: 427–35.
18. **Cech TR.** 2000. The ribosome is a ribozyme. *Science* **289**: 878–9.
19. **Voorhees RM**, **Weixlbaumer A**, **Loakes D**, **Kelley AC**, *et al*. 2009. Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome. *Nat Struct Mol Biol* **16**: 528–33.
20. **Ehrenberg M**, **Kurland CG.** 1984. Costs of accuracy determined by a maximal growth rate constraint. *Q Rev Biophys* **17**: 45–82.
21. **Kurland CG.** 1992. Translational accuracy and the fitness of bacteria. *Annu Rev Genet* **26**: 29–50.
22. **Valentine DL.** 2007. Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nat Rev Microbiol* **5**: 316–23.
23. **Lecompte O**, **Ripp R**, **Thierry J-C**, **Moras D**, *et al*. 2002. Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. *Nucleic Acid Res* **30**: 5382–90.
24. **Poole A**, **Jeffares D**, **Penny D.** 1999. Early evolution: prokaryotes, the new kids on the block. *BioEssays* **21**: 880–9.
25. **Brinkman H**, **Phillipe H.** 1999. Archaea sister-group of bacteria? Indications from tree reconstruction artifacts in ancient phylogenies. *Mol Biol Evol* **16**: 817–25.
26. **Kurland CG**, **Collins LJ**, **Penny D.** 2006. Genomics and the irreducible nature of eukaryotic cells. *Science* **312**: 1011–14.
27. **Kurland CG**, **Canback B**, **Berg OG.** 2007. The origins of modern proteins. *Biochimie* **89**: 1454–63.
28. **Kurland CG.** 1974. Functional organization of the 30S ribosome subunit. In **Nomura M**, **Tissiere A**, **Lengyel P.** ed; *Ribosomes*. Woodbury, NY: Cold Spring Harbor Laboratory Press. p 309–31.
29. **Liljas A.** 2004. *Structural Aspects of Protein Synthesis*. Singapore: World Scientific Publishing Co.Pte. Ltd.
30. **Cairns-Smith AG.** 1985. *Seven Clues to the Origin of Life*. Cambridge: University Press.
31. **Kauffman SA.** 1993. *The Origins of Order*. New York: Oxford University Press.
32. **Egel R.** 2009. Peptide-dominated membranes preceding the genetic takeover by RNA: latest thinking on a classic controversy. *BioEssays* **31**: 1100–9.
33. **Goldberg AL.** 1972. Degradation of abnormal proteins in Escherichia coli. *Proc Natl Acad Sci* **69**: 422–36.
34. **Goldberg AL.** 2003. Protein degradation and protection against unfolded or damaged proteins. *Nature* **426**: 895–9.
35. **Prakash S**, **Matouschek A.** 2004. Protein unfolding in the cell. *Trends Biochem Sci* **29**: 593–600.
36. **Oliveberg M**, **Wolynes PG.** 2006. The experimental survey of protein-folding energy landscapes. *Q Rev Biophys* **38**: 275–88.
37. **Tyndall JD**, **Nall T**, **Fairlie DP.** 2005. Proteases universally recognize beta strands in their active sites. *Chem Rev* **105**: 973–99.
38. **Ellis RJ.** 2001. Macromolecular crowding: obvious but underappreciated. *Trends Biochem Sci* **26**: 597–604.
39. **Cookson MR.** 2005. The biochemistry of Parkinson's disease. *Annu Rev Biochem* **4**: 29–52.
40. **Chitil F**, **Dobson CM.** 2006. Protein misfolding, functional amyloid and human disease. *Annu Rev Biochem* **75**: 333–66.
41. **Gidalevitz T**, **Ben-Zvi A**, **Ho KH**, **Brignull HR**, *et al*. 2006. Progressive disruption of cellular protein folding in models of polyglutamine diseases. *Science* **246**: 1471–84.
42. **Powers ET**, **Morimoto RI**, **Dillin A**, **Kelly JW**, *et al*. 2009. Biological and chemical approaches to diseases of proteostasis deficiency. *Annu Rev Biochem* **78**: 959–91.
43. **Schubert U**, **Antón LC**, **Gibbs J**, **Norburg CC**, *et al*. 2000. Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature* **404**: 770–4.
44. **Pratt JM**, **Petty J**, **Riba-Garcia I**, **Robertson DHL**, *et al*. 2002. Dynamics of protein turnover, a missing dimension in proteomics. *Mol Cell Proteomics* **1**: 579–91.
45. **Kurland CG**, **Berg OG.** 2010. A hitchhiker's guide to evolving networks. In **Caetano-Anolles G.** ed; *Evolutionary Genomics and Systems Biology*. Hoboken, New Jersey: Wiley and Sons. p 363-96.
46. **Doolittle RF.** 1995. The multiplicity of domains in proteins. *Annu Rev Biochem* **64**: 287–314.
47. **Hancock REW**, **Chapple DS.** 1999. Peptide antibiotics. *Antimicrob Agents Chemother* **43**: 1317–23.
48. **Finking R**, **Marahiel MA.** 2004. Biosynthesis of non-ribosomal peptide. *Annu Rev Microbiol* **58**: 453–88.
49. **Marahiel MA.** 2009. Working outside the protein-synthesis rules: insights into non-ribosomal peptide synthesis. *J Pept Sci* **15**: 607–799.
50. **Keefe AD**, **Szostak JW.** 2001. Functional proteins from a random-sequence library. *Nature* **410**: 715–8.
51. **Szostak JW.** 2009. Systems chemistry on early earth. *Nature* **459**: 171–2.
52. **Powner MW**, **Gerland B**, **Sutherland JD.** 2009. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459**: 239–242.
53. **Freistroffer DV**, **Kwiatkowski M**, **Buckingham RH**, **Ehrenberg M.** 2000. The accuracy of codon recognition by polypeptide release factors. *Proc Natl Acad Sci* **97**: 2046–51.
54. **Zhang S**, **Marini DM**, **Hwang W**, **Santoso S.** 2002. Design of nanostructured biological materials through self-assembly of peptides and proteins. *Curr Opin Chem Biol* **6**: 865–71.
55. **Yokoi H**, **Kinoshita T**, **Zhang S.** 2005. Dynamic reassembly of peptide RADA16 nanofiber scaffold. *Proc Natl Acad Sci* **102**: 8414–9.
56. **Szostak JW**, **Bartel DP**, **Luisi PL.** 2001. Synthesizing life. *Nature* **409**: 387–90.
57. **Chen IA**, **Roberts RW**, **Szostak JW.** 2004. The emergence of competition between model protocells. *Science* **305**: 1474–6.
58. **Hanczye MM**, **Szostak JW.** 2004. Replicating vesicles as models of primitive cell growth and division. *Curr Opin Chem Biol* **8**: 660–4.
59. **Tooze J.** 1983. *DNA Makes RNA Makes Protein*. Amsterdam: Elsevier.
60. **Caetano-Anolles D**, **Kim KM**, **Mittenthal J**, **Caetano-Anolles G.** 2010. Proteome evolution and the metabolic origins of translation and cellular life. *J Mol Evol* in press.