J Mol Evol (2000) 51:416–421
DOI: 10.1007/s002390010104

JOURNAL OF MOLECULAR EVOLUTION

Letter to the Editor

RNA Folding Argues Against a Hot-Start Origin of Life

Vincent Moulton,[1,*] Paul P. Gardner,[2] Robert F. Pointon,[3] Lawrence K. Creamer,[4] Geoffrey B. Jameson,[2] David Penny[1]

[1] Institute of Molecular BioSciences, Massey University, PO Box 11-222, Palmerston North, New Zealand
[2] Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand
[3] Institute of Information Sciences and Technology, Massey University, Palmerston North, New Zealand
[4] Food Science Section, Dairy Research Institute, Palmerston North, New Zealand

Received: 1 March 2000 / Accepted: 14 June 2000

Abstract. Opinion is strongly divided on whether life arose on earth under hot or cold conditions, the hot-start and cold-start scenarios, respectively. The origin of life close to deep thermal vents appears as the majority opinion among biologists, but there is considerable biochemical evidence that high temperatures are incompatible with an RNA world. To be functional, RNA has to fold into a three-dimensional structure. We report both theoretical and experimental results on RNA folding and show that (as expected) hot conditions strongly reduce RNA folding. The theoretical results come from energy-minimization calculations of the average extent of folding of RNA, mainly from 0–90°C, for both random sequences and tRNA sequences. The experimental results are from circular-dichroism measurements of tRNA over a similar range of temperatures. The quantitative agreement between calculations and experiment is remarkable, even to the shape of the curves indicating the cooperative nature of RNA folding and unfolding. These results provide additional evidence for a lower temperature stage being necessary in the origin of life.

Key words: Circular dichroism — Origin of life — RNA folding — RNA secondary structure — RNA world

* *Present address:* Physics and Mathematics Dept., Mid-Sweden University, S 851-70 Sundsvall, Sweden
Correspondence to: D. Penny; *email*: d.penny@massey.ac.nz

Opinion is strongly divided over whether life arose on earth under hot or cold conditions, the hot-start and cold-start scenarios, respectively. The suggested conditions range from >85°C in deep-sea thermal vents (Russell and Hall 1988; Pace 1991) to water under high pressures at the bottom of a frozen ocean (Bada et al. 1994). Though popular opinion still appears to favors the hot-start scenario, "a significant number of holdouts" (*Science* 283: 156–157 [1999]) are providing strong evidence that lower temperatures are much more favorable for the origin of life (see also Miller and Bada 1988).

Two approaches are commonly used for understanding of the origin of life: working forward from simpler chemical systems and working backward from existing organisms. The main support for a hot-start origin of life has come from the second approach by inferring properties of the earliest forms of modern life. For example, the rooting of the tree of life using duplicated genes (Gogarten et al. 1989; Iwabe et al. 1989) appeared to support the last universal common ancestor (LUCA) being hyperthermophilic. This, together with the discovery of living organisms at thermal vents (black smokers) in the deep ocean, lent credibility and drama to the concept.

However, with improved methods of phylogenetic analysis—in particular identifying the slowest evolving sites in a sequence—the root no longer fell among hyperthermophiles but among mesophilic organisms (Brinkman and Philippe 1999; Lopez et al. 1999). Simi-

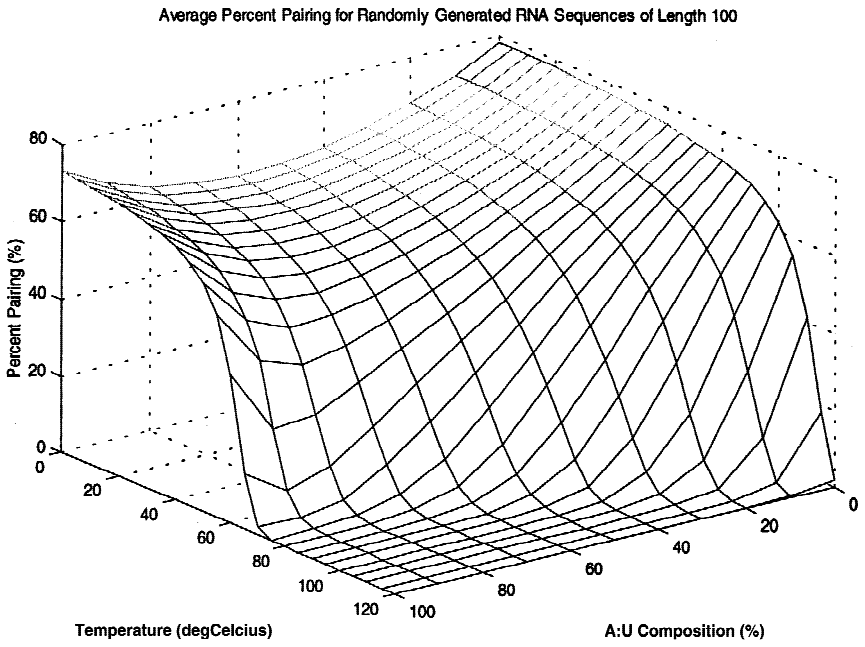

Fig. 1. Percentage of nucleotide pairing and extent of energy minimization, by temperature and nucleotide composition. Results are averaged over 1000 runs for random sequences of length 100. The optimal (minimum free-energy) structures were found by the Vienna package (Hofacker et al. 1994). Folding of RNA decreases rapidly above about 30°C.

larly, a recent study that projected nucleotide composition back in time (Galtier et al. 1999) predicted that the G+C content of the ancient genomes were characteristic of mesophilic rather than thermophilic organisms. This means that at present there is no firm evidence from biology for a hyperthermophilic LUCA, and so it leaves unresolved the temperature optima of earlier forms of life.

We thus turn to evidence from the second (working forward) approach, in particular the consideration of intermediate stages in the origin of life prior to the LUCA. An RNA world is an obligatory stage in current models of the origin of life (Gesteland et al. 1999), a stage when RNA served as both the primary catalyst and stored the genetic information. However, it is already well known that both RNA (Lindahl 1967) and its constituent nucleotides (Levy and Miller 1998) have short half-lives at hyperthermophilic temperatures. For example, RNA bases can have a half-life as short as 19 days at 100°C for cytosine (Levy and Miller 1998). Similarly, an RNA molecule is chemically unstable to the extent that an RNA molecule 2000 nucleotides long is expected to have one break every 26 s at 100°C (Lindahl 1967).

In addition to the problems of thermolability of RNA and its nucleotides, there is the third problem, the extent of folding at high temperatures. To be functional as a catalyst, RNA has to fold into a three-dimensional structure, and it is expected that the extent of folding will decrease at higher temperatures. In addition, an RNA molecule can fold into an extremely large number of structures of similar free energy (Uhlenbeck 1995), and so specificity of structure is not expected to be high. When specificity is not strongly delineated, a small free-energy difference (ΔG) in either activation energy (kinetics) or thermodynamic stability is accentuated at lower temperatures; ΔG becomes relatively greater with respect to thermal energy (kT). Conversely, at higher temperatures, the folding that does occur is expected to be less specific and to interchange more quickly among structures of similar free energy.

There has been little recent experimental work on the folding of RNA into secondary and tertiary structures as a function of temperature, and the earlier work was not oriented to origin-of-life hypotheses. Thus, to study the effect of high temperatures on RNA structure we carried out a combined theoretical and experimental study. RNA secondary structures were generated by standard free-energy minimization calculations using the programs M-fold (Zuker 1989; Matthews et al. 1999) and the VIENNA package (Hofacker et al. 1994). This was done for ensembles of random sequences, as well as for tRNA sequences of *Escherichia coli*, yeast (*Saccharomyces cerevisiae*), and cow (*Bos taurus*). We then made circular-dichroism measurements on tRNA at different temperatures to measure directly the amount of secondary (helical) structure as a function of temperature.

We report that both theoretical and experimental results on RNA folding as a function of temperature show that hot conditions strongly reduce the extent of folding. Figure 1 shows the percentage of nucleotide pairing for 1000 random sequences, each 100 nucleotides long. The results are plotted at 5° intervals from 0 to 120°C, and for nucleotide composition from 0–100% AU. Results for lengths over the range 48–150 nucleotides were very similar (data not shown). At lower temperatures (say, 0–35°C) we see high pairing for the optimal structures irrespective of nucleotide composition. The extent of base-pairing decreases fairly rapidly above about 35°C, until there is no pairing for 100% AU above about 70°C. At low temperature there is slightly less pairing around

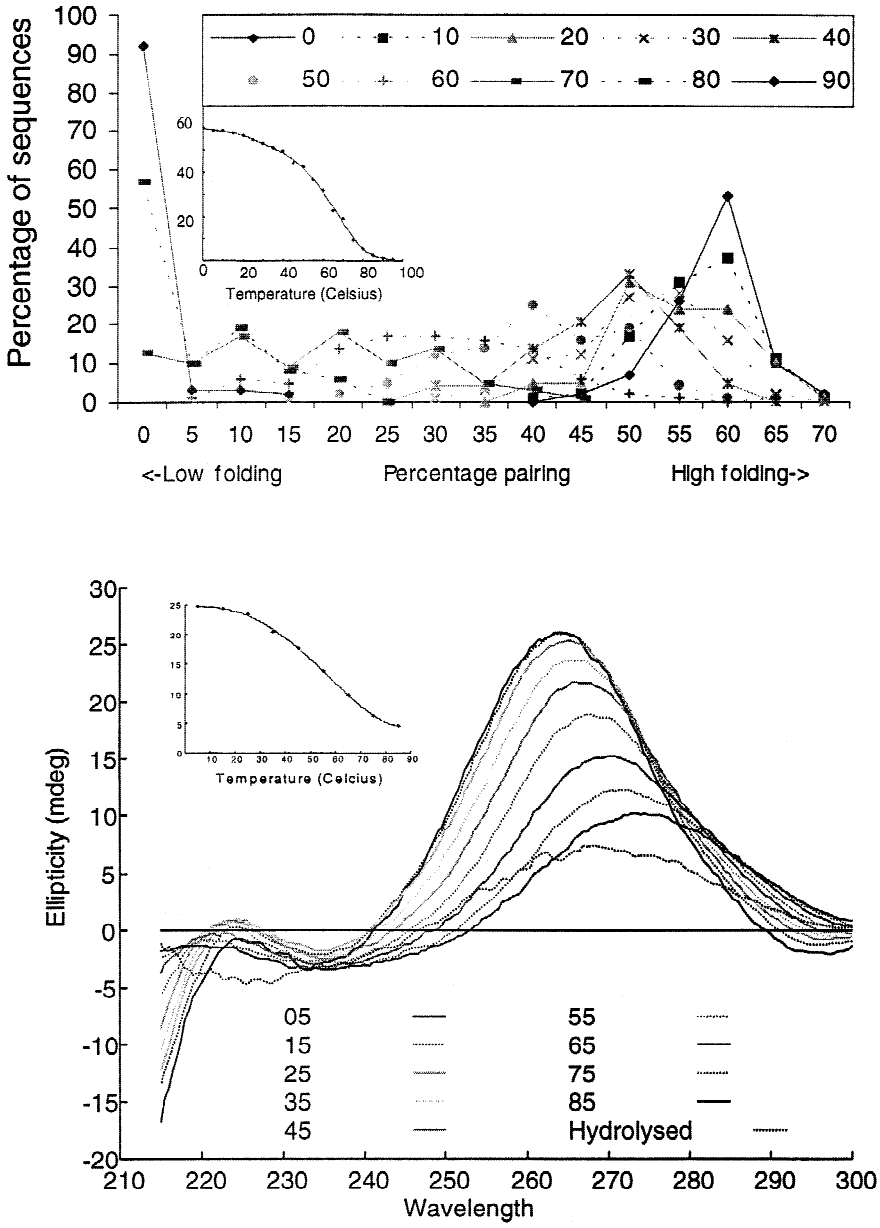

Fig. 2. Extent of RNA folding as a function of temperature for yeast (*S. cerevisiae*) tRNAs. (upper) The distribution function (for 5% intervals) for the percentage nucleotide pairing as calculated by Mfold (Matthews et al. 1999). It is averaged over 20 tRNA sequences with unmodified nucleotides and is shown for temperatures spanning the range 0–90°C. The S-shaped curve of the upper insert shows the average percent pairing of RNA as a function of temperature, indicates the cooperative nature of RNA folding and unfolding. (lower) Circular dichroism in the range 210–320 nm of the *S. cerevisiae* tRNA mixture (Sigma, catalog R-9001) plotted at 10° intervals from 5–85°C. Measurements are on twice-dialyzed tRNA in 0.010 *M* phosphate buffer at pH 7.0, measured on a Jasco J720 CD spectrometer. The spectral changes observed are reversible over this temperature range. The CD of a completely hydrolyzed sample is shown for reference (no folding). The lower insert shows the signal at 260 nm versus temperature, indicating an approximate melting temperature of 62°C for the ensemble of tRNAs. tRNAs from *E. coli* and cow show similar behavior and have melting temperatures of 58°C and 68°C.

50% AU composition, but at higher temperatures the amount of pairing increases with increasing GC content. The higher pairing of higher GC content at higher temperatures is expected, but by itself is not helpful for an early RNA world where, if anything, AU is thought to have predominated. Adenine is much easier to synthesize under prebiotic conditions (Levy et al. 1998), and ribozyme activity is strongly reduced without the full complement of four nucleotides (Rogers and Joyce 1999). In addition, cytosine is especially unstable at high temperatures (Levy and Miller 1998). For these reasons we were particularly interested in the folding of sequences with low to average AU content, and they are predicted not to fold at the higher temperatures.

These first results are the average pairing for 1000 sequences. This is informative because although some specific sequences may have good pairing at higher temperatures, it is difficult to explain the origin of life under a specific set of conditions if most sequences did not even fold. Nevertheless, it is useful to know the variability in pairing between sequences, just in case a small percentage of sequences folded well at high temperature. We therefore computed the distribution of folding for both randomly generated sequences and tRNAs.

The predicted percentage pairing for yeast tRNAs is shown in Fig. 2 (upper) for different temperatures (the tRNA sequences used in calculations had only unmodified nucleotides). The results for average pairing are similar to the random sequences in Fig. 1. High nucleotide pairing (and thus folding) is found only at lower temperatures, low or no pairing occurs at the temperatures expected of thermophiles. The distribution of the percent pairing at temperatures in the range 0–90°C is as expected and does not have a high tail of strong pairing

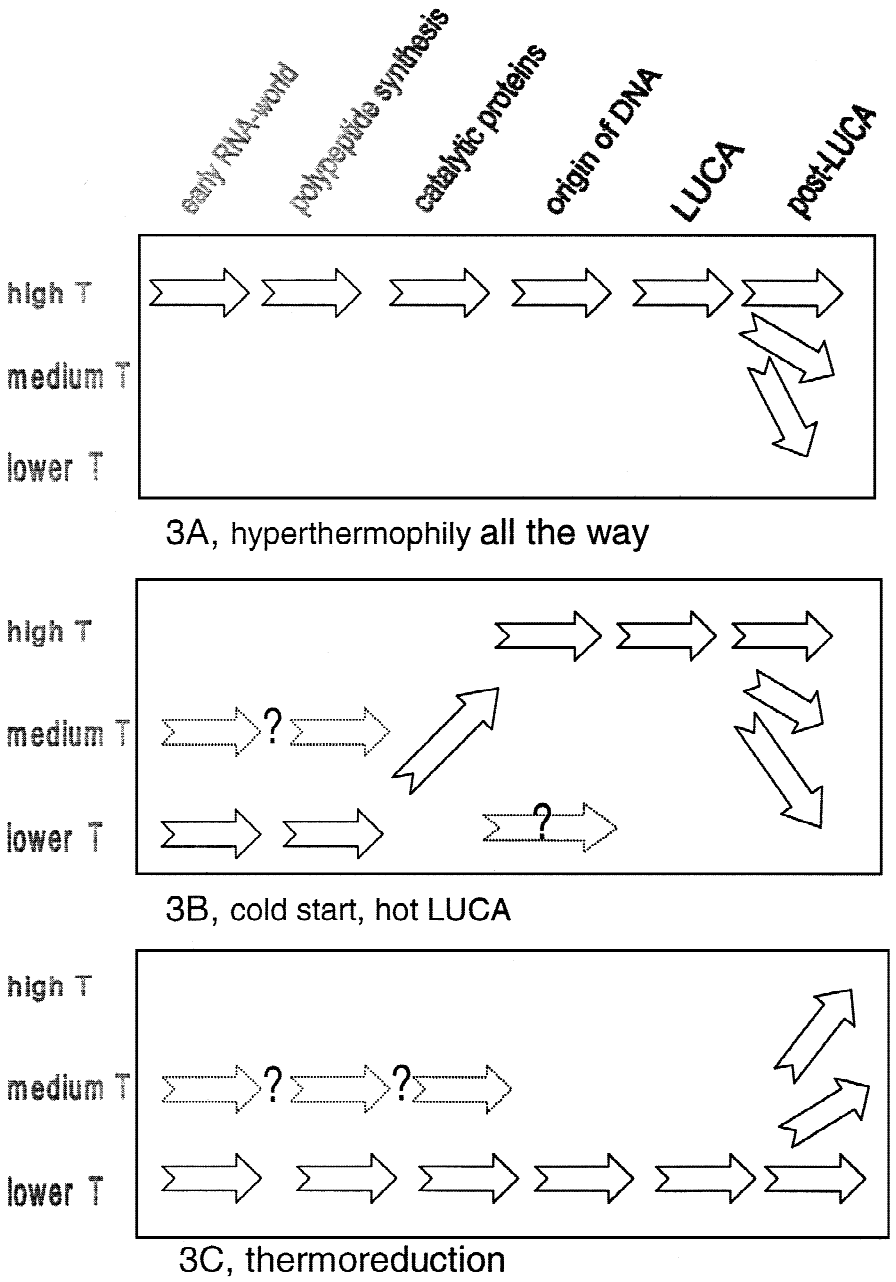

Fig. 3. Three scenarios for the phylogenetic position of extreme thermophiles. **A,** all later stages of the origin of life, from the early RNA world to the last common ancestor (LUCA), occur at high temperature (>70°C). **B,** the RNA world occurs at lower temperatures (0–30°C), but, after the advent of proteins and DNA, life adapted to high temperatures. Under this scenario, the LUCA was thermophilic. Although the RNA world appears more likely at lower temperatures, a later adaptation to medium temperature (30–70°C) cannot be excluded and this option is shown as dotted arrows. **C,** the thermoreduction hypothesis. All stages of the origin of life occur at mesophilic temperatures and thermophiles only arose after LUCA (Forterre 1996). All biochemical evidence on RNA thermolability, including the present results, excludes option A. Although we favor option C, B cannot yet be eliminated.

at high temperatures. Only at lower temperatures does the fraction of unfolded sequences become small. We do not find it likely that a small fraction of sequences would fold well at higher temperatures. There is an additional conclusion. In contrast to tRNAs from hyperthermophiles (Yokoyama et al. 1987; Kowalak et al. 1994), the RNA structures of mesophiles do not enjoy enhanced stability relative to a random sequence.

The calculations that predict a low degree of folding at high temperatures were then tested experimentally. Circular dichroism (CD) spectra, in the presence and absence of 200 μM $MgCl_2$, of tRNAs (thus with some modified nucleotides; see Auffinger and Westhof 1998) from *E. coli,* yeast, and cow were each recorded over the temperature range 5–85°C. For each species, a mixture of tRNAs was used in order to measure average behavior and thus avoid any single, anomalous tRNA. The height of the peak in the 270-nm region (or depth of the valley around 210 nm) is a measure of the amount of secondary and tertiary structure (Hilbers et al. 1976; Privalov and Filimonov 1978). Again there is little folding (and therefore structure) at the highest temperatures. The loss of structure is reversible if the temperature is lowered. The addition of Mg^{2+} does delay the onset of the loss of secondary and tertiary structure. This is in substantial agreement with earlier work on a single yeast tRNA (tRNAMet) (Stein and Crowther 1976 and references therein), where, by UV-visible spectroscopy, mM concentrations of Mg^{2+} delay the onset of loss of secondary structure from ~35°C to ~65°C. In other work, the pres-

ence of Mg^{2+} was observed to accelerate RNA breakdown (Lindahl 1967) and to increase the rate of rearrangement of tertiary structures (Chetverina et al. 1999). For the purpose of comparison with the theoretical calculations, only the results of experiments conducted in the absence of Mg^{2+} are shown in the lower panel of Fig. 2.

The quantitative agreement between calculations and experiment is remarkable, even to the shape of the curves (inserts in Fig. 2, upper and lower). Theoretical calculations did not consider the effects of modified nucleotides that may enhance local flexibility of psychrophiles (Dalluge et al. 1997) and enhance global stability for hyperthermophiles (Kowalak et al. 1994). In addition, the folding programs embed the molecule in an isotropic environment wherein many suboptimal secondary structures of essentially identical free energy are generated. There is evidence (Wu and Tinoco 1998) for uniqueness of structure for folded RNAs in solution, where specific solvent-solute and cation-phosphate interactions are present to discriminate among structures of equal energy (isoergonic). The NMR evidence also indicates that for a given set of conditions, even for relatively short fragments of less than 54 nucleotides, an essentially unique secondary structure is observed. In addition, interconversion between structures when conditions are altered may occur at cool temperatures (25°C) (Brion and Westhof 1997; Wu and Tinoco 1998). By concentrating on the amount of folding, rather than studying the whole landscape of suboptimal structures (see Moulton et al. 2000), both calculation and measurement indicate low RNA folding at high temperatures.

These results provide additional support for a lower temperature stage in the origin of life. Not surprisingly, in vitro evolution experiments with RNA (Wright and Joyce 1997) are carried out at mesophilic temperatures, not those of extreme thermophiles. The results, theoretical and experimental, are strong evidence against an RNA world at hyperthermophilic temperatures. There are, however, many stages in the origin of life, and the results do not necessarily exclude other stages occurring at high temperatures. The results reported here are based on modern biochemistry but do not allow us to infer the very earliest stages of the origin of life. But for the later stages, based on modern biochemistry, we can classify stages as an early RNA world, the origin of polypeptide synthesis, then catalytic proteins, DNA, and finally LUCA (Penny and Poole 1999). Figure 3 shows three alternative hypotheses for the origin of extreme thermophiles. In 3A all later stages are hyperthermophilic, in 3B thermophily arises after the RNA world, and in 3C thermophily arises by thermoreduction more recently than the LUCA. Alternatively, given that folded functional RNA molecules require mesophilic, if not psychrophilic conditions, a hyperthermophilic stage in the origin of life requires then that small peptide fragments or other moieties are present to stabilize folded functional forms of RNA. There is little evidence at this stage for such species.

Together with the thermolability results of Lindahl (1967) and Levy and Miller (1998), the present results on RNA folding are a third strike against an early hyperthermophilic RNA world option (3A). The present results show no secondary structure at high temperatures, especially with the high AU content expected at early stages. Although we favor option 3C, the present results, together with those of Lindahl (1967) and Levy and Miller (1998) really only disproves alternative 3A.

Thus, irrespective of the very earliest stages of the origin of life, there is converging evidence that the study of the RNA world to the last universal common ancestor of all life should focus on more mesophilic conditions. Within the framework of an RNA-based predecessor to the protein/DNA life forms that now infest the earth, thermophily for such a predecessor is not supported.

References

Auffinger P, Westhof E (1998) Location and distribution of modified nucleotides in tRNA. *In:* Grosjean H, Benne R (eds) Modification and editing of RNA. Washington, DC: ASM Press

Bada JL, Bigham C, Miller SL (1994) Impact melting of frozen oceans on the early earth: implications for the origin of life. Proc Natl Acad Sci USA 91:1248–1250

Brinkmann H, Philippe H (1999) Archaea sister-group of Bacteria? Indications from tree reconstruction artifacts in ancient phylogenies. Mol Biol Evol 16:817–825

Brion P, Westhof E (1997) Hierarchy and dynamics of RNA folding. Ann Rev Biophys Biomol Struct 26:113–137

Chetverina HV, Demidenko AA, Ugarov VI, Chetverin AB (1999) Spontaneous rearrangements in RNA sequences. FEBS Lett 450: 89–94

Dalluge JJ, Hamamoto T, Horikoshi K, Morita RY, Stetter DO, McCloskey JA (1997) Posttranscriptional modification of tRNA in psychrophilic bacteria. J Bacteriol 179:1918–1923

Forterre P (1996) A hot topic: the origin of hyperthermophiles. Cell 85:789–792

Galtier N, Tourasse N, Gouy M (1999) A nonhyperthermophilic common ancestor to extant life forms. Science 283:220–221

Gesteland RF, Cech TR, Atkins JF (eds) (1999) The RNA world, 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Lab Press

Gogarten JP, Kibak H, Dittrich P, Taiz L, Bowman EJ, Bowman BJ, Manolson MF, Poole RJ, Date T, Oshima T, Konishi J, Denda K, Yoshida M (1989) Evolution of the vacuolar H^+-ATPase: implications for the origin of eukaryotes. Proc Natl Acad Sci USA 86: 6661–6665.

Hilbers CW, Robillard GT, Shulman RG, Blake RD, Webb PK, Fresco R, Riesener D (1976) Thermal unfolding of yeast glycine transfer RNA. Biochemistry 15:1874–1882

Hofacker I, Fontana W, Stadler P, Bonhoeffer L, Tacker M, Schuster P (1994) Fast folding and comparison of RNA secondary structures. Monat Chem 125:167–188

Iwabe N, Kuma K-I, Hasegawa M, Osawa S, Miyata T (1989) Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc Natl Acad Sci USA 86:9355–9359

Kowalak JA, Dalluge JJ, McCloskey JA, Stetter KO (1994) The role of posttranscriptional modification in stabilization of transfer RNA from hyperthermophiles. Biochemistry 33:7869–7876

Levy M, Miller SL (1998) The stability of the RNA bases: implications for the origin of life. Proc Natl Acad Sci USA 95:7933–7938

Levy M, Miller SL, Oró J (1998) Production of guanine from NH_4CN polymerizations. J Mol Evol 49:165–168

Lindahl T (1967) Irreversible heat inactivation of transfer ribonucleic acids. J Biol Chem 242:1970–1973

Lopez P, Forterre P, Philippe H (1999) The root of the tree of life in light of the covarion model. J Mol Evol 49:496–508

Mathews DH, Sabina J, Zuker M, Turner DH (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J Mol Evol 288:911–940

Miller SL, Bada JL (1988) Submarine hot springs and the origin of life. Nature 334:609–611

Moulton V, Zuker M, Steel MA, Pointon R, Penny D (2000) Metrics on RNA secondary structure. J Comput Biol 7:277–292

Pace N (1991) Origin of life—facing up to the physical setting. Cell 65:531–533

Penny D, Poole AM (1999) The nature of the universal ancestor. Curr Opin Genet Devel 9:672–677

Privalov PL, Filimonov VV (1978) Thermodynamic analysis of transfer RNA unfolding. J Mol Biol 122:447–464

Rogers J, Joyce GF (1999) A ribozyme that lacks cytidine. Nature 402:323–325

Russell MJ, Hall AJ (1988) Submarine hot springs and the origin of life. Nature 336:117

Stein A, Crowther DM (1976) Equilibrium binding of magnesium (II) by *Escherichia coli.* Biochemistry 15:157–160

Uhlenbeck OC (1995) Keeping RNA happy. RNA 1:4–6

Wright MC, Joyce GF (1997) Continuous in vitro evolution of catalytic function. Science 276:614–617

Wu M, Tinoco I Jr (1998) RNA folding causes secondary structure rearrangement. Proc Natl Acad Sci USA 95:11555–11560

Yokoyama S, Watanabe K, Miyazawa T (1987) Dynamic structures and functions of transfer ribonucleic acids from extreme thermophiles. Adv Biophys 23:115–147

Zuker M (1989) On finding all suboptimal foldings of an RNA molecule. Science 244:48–52