ORIGIN OF TRANSLATION: THE HYPOTHESIS OF PERMANENTLY ATTACHED ADAPTORS

SANJAY TYAGI

School of Environmental Sciences Jawaharlal Nehru University New Delhi – 110 067, India

(Received 25 February, 1981; in revised form 4 August 1981)

Abstract. A mechanism for prebiotic translation is proposed in which primeval transfer-RNA (adaptors) are assumed to be permanently associated with messenger nucleic acid molecules. Residual 'fossil' evidences are found to be present within the base sequences of contemporary tRNAs, suggesting the existence of inter-primal-tRNA interactions necessary for the mechanism. The structure of proposed primal-tRNA is such that it can not only choose its own amino acid in the absence of aminoacyl synthetase, but can also associate nonspecifically with adjacent primal-tRNA molecules attached to the neighbouring codons. Such associations can give rise, through cooperative binding between message and adaptors to the 'static template surfaces' which can direct translation of nucleotide sequences into those of amino acids. The origins of ribosomes and contemporary genetic code are suggested by this hypothesis. Proposed structures and processes are thermodynamically compatible. The approximate date of occurence of the proposed system is calculated, which is consistent with the period of occurence of the earliest organisms with ribosomes.

Key words: Translation/Genetic code/Ribosomes/Static template surfaces/Adaptors/Primal-tRNA.

1. Introduction

In experiments aimed at tracing the origins of life on Earth by simulating probable pre-biotic conditions, the occurrence of significant amounts of precursors of biomolecules, including basic building blocks (Miller and Orgel, 1974) as well as their activated species has been demonstrated. For example, polyphosphates of nucleotides (Sagan and Ponnamperuma, 1973) and adenylates of amino acids (Paecht-Horowitz, 1976) have been synthesized under such primitive conditions. For the polymerization of monomeric units, a necessary condition for any form of life, several schemes have been proposed. Polymerization of amino acids into proteins has been suggested as having occurred by chemical condensation (Steinman *et al.*, 1965; Hiroyuki and Egami, 1979), thermal synthesis (Fox, 1960) or by synthesis on clay surfaces (Paecht-Horowitz, 1974; Rao *et al.*, 1980). The clay synthesis is particularly attractive because this fulfils most of the requirements of prebiotic polymerization, i.e., the random condensation with right kinds of bonds, formation of sufficiently long polymers and occurrence close to a water reservoir (Rao *et al.*, 1980). Unfortunately, no satisfactory scheme is available for the polymerization of nucleic acids; chemical synthesis either gives rise to the wrong type of bonds or is implausible under prebiotic conditions. If we consider the profuse polymerization carried out by clay minerals (according to Bernal, 1951; proteinaceous material three times the total existing biotic proteins is entrapped in the clay deposits of the Earth), we can speculate a plausible mechanism of nucleotide polymerization which can produce sufficient amounts of random polynucleotides.

Origins of Life **11** (1981) 343–351. 0302–1688/81/0114–0343 $ 01.35.

A great amount of polymerized amino acids might have been leaching every day into the ocean from clay beds. These random polypeptides would possibly have a variety of weak catalytic activities. The occurrence of a 'nucleotidyl polymerase' activity (which polymerises nucleotides) in such a mixture is quite probable if we consider the strikingly small size of a contemporary enzyme, 'the terminal nucleotidyl polymerase' of the thymus (Kornberg, 1974), which polymerizes nucleotides without any template. The polypeptides leaching out from clays may form 'microspheres' in water (Fox and Dose, 1972) owing to the physical conditions prevailing there. There is a fair probability of some such hypothetical nucleotidyl polymerase molecules being trapped inside some of the microspheres. If this happens, a large number of polynucleotides would be manufactured in these microspheres. Such a condition has been simulated experimentally and detectable amounts of polynucleotides have been made (Liebl *et al.*, 1978). Finally, for the self-replicating system to originate, replication of the polynucleotides is required; this can not occur with sufficient fidelity and efficiency in case of longer polymers, without the help of an intrinsic catalyst (Eigen and Schuster, 1979). If the system is to survive natural decay, it has to produce its own 'template dependent polymerase'. Most of the microspheres would die out soon except those which have polynucleotides and mechanisms capable of producing template dependent polymerases. How these polynucleotides could have been translated into the appropriate sequences of amino acids is the subject of this paper. It is not obligatory for the proposed mechanism of prebiotic translation to assume the above scheme for the origin of polynucleotides.

2. Mechanism of Primeval Translation

In prebiotic conditions, when no ribosomes and aminoacyl synthetases were present, translation was required to be carried out by a mechanism, which could function in their absence with satisfactory fidelity and speed. The structure of prebiotic transfer-RNA (primal-tRNA) had to differ from that of contemporary ones because no enzyme existed to charge them and no ribosome existed to assist them in translation. A model for the structure of primal-tRNA has been proposed by Hopfield (Hopfield, 1978) with sufficient evidence in support, according to which the acceptor system (bearing the amino acid), instead of being directed upward, away from the main body, was situated near the anticodonic region. Such a condition can be achieved by a different type of folding of tRNA in which the 3′ acceptor stem is brought close to the anticodonic stretch keeping the $T\Psi C$ stem intact (Figure 1(b)). This brings the amino acid into the immediate vicinity of the anticodon. The sequence of primal-tRNA is assumed to have differed considerably from modern tRNA so that the proposed secondary structure is more stable than the modern one. The close interaction between amino acid and its anticodon is desired because this can lead to recognization of a codon by an amino acid. Several physical properties (for example, hydrophobicity, polarity and bulkyness etc.) of amino acids show correlations with their respective *anticodonic* dinucleotides (Weber and Lacey, 1978; Jungck, 1978). Some amount

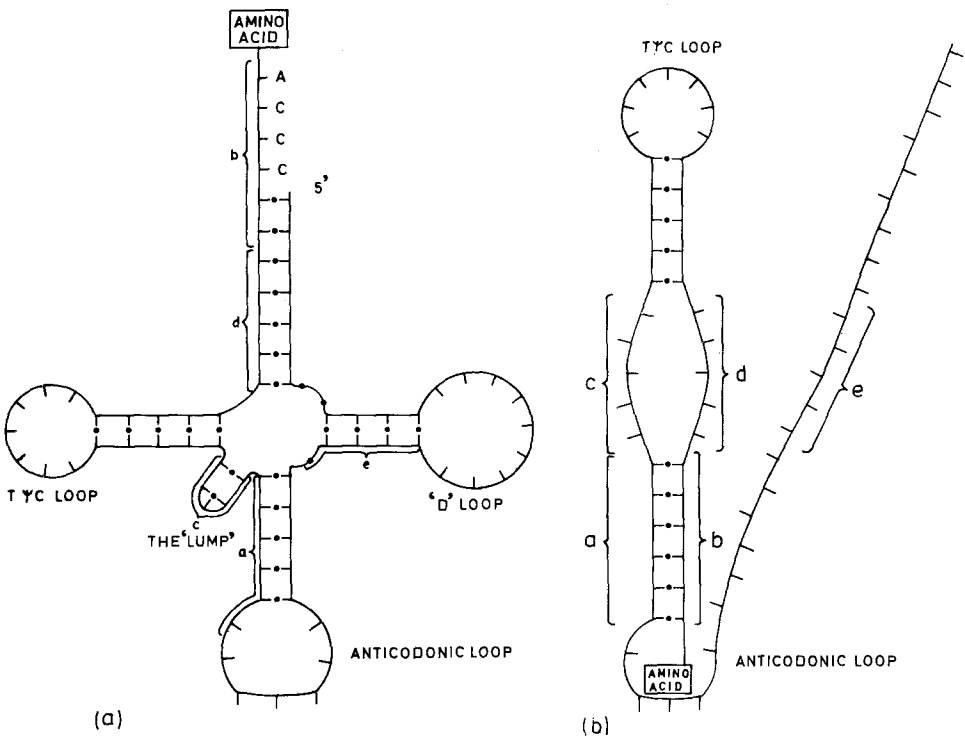

Fig. 1. Secondary structure of contemporary and primal-tRNAs (a) contemporary tRNA; (b) primal-tRNA as proposed by Hopfield. Its primary sequence differed from modern one in such a way that proposed secondary folding was more stable.

of specificites also have been shown to exist between several amino acids and homo-polynucleotides which contain the respective anticodons (Lacey *et al.*, 1979). The conformation of that face of the anticodon which does not interact with the codon is such that it can accomodate an amino acid molecule with a good fit (Balasubramanian *et al.*, 1980). The close proximity of amino acid with its anticodon can help it in choosing its codon and in discriminating from the wrong one.

When sequences of contemporary tRNAs (of *E. coli*) were refolded according to the proposed pattern, it was observed that on the average, 2.5 out of 6 bases, of the region a and b (Figure 1(b)) are paired. The region a and b are located at different positions in contemporary t-RNAs (Figure 1(a)), so that they do not have any association with each other. Then why do they pair (the pairing is statistically significant)? Hopfield (1978) suggests that primal-tRNA had such a correspondence between regions a and b that all of the bases of the two regions were complementary; the observed pairing between said regions of contemporary tRNA is reminiscent of that archaic correspondence.

While refolding modern tRNA to resemble the primal one, it was observed that an 'internal loop' is inevitably formed in c and d regions (Figure 1(b)) in almost all of the t-RNAs of *E. coli* and those of *Bacillus subtillis* and bases of region c are never

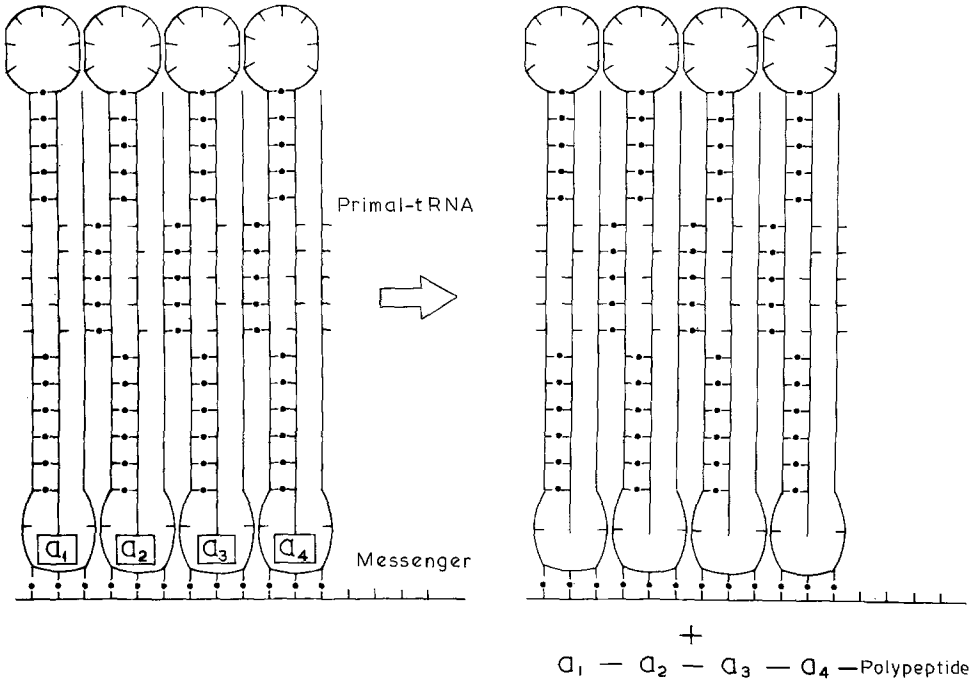

Fig. 2. Inter primal-tRNA association leading to permanent adaptor messenger complex. a_1, a_2, a_3, ... represent amino acids. A polypeptide is formed by virtue of close proximity.

complementary to those of d. However, interestingly, the bases of region c were found pairing with those of region e (Figure 2) of not only the same tRNA but of all other tRNAs also, such that any two adjacent tRNAs could associate with each other through region c and e. On the basis of this observation, it is hypothesized that primal-tRNA could not only interact with messenger nucleic acid (through anticodon), but with neighbouring primal-tRNA also (c of one with e of the next). Interprimal-tRNA association would lead to *all-or none limit* type cooperative binding (according to which the macromolecule is either entirely covered with ligands, without gaps or remains utterly bare) with messenger, resulting in formation of a more-or-less permanent complex.

Now this complex can function as a template for amino acids. It contains readymade sites for amino acids (or rather their active species) which can place themselves on anticodons. The amino acid adenylates are the most appropriate candidates for these sites. They have a high energy bond which can be utilized for the formation of an aminoacyl linkage between the acceptor stem and the carboxyl group. An acylated amino acid in slightly basic (say pH 8) medium is neutral but still activated. If two such amino acid molecules are aligned properly and held closely for sufficiently long time, the peptide bond would be formed. The mechanism would be most probably similar to the clay induced condensation (Paecht-Horowitz, 1976).

Small amount of specificity between amino acid anticodon can lead to high fidelity

translation if 'time delay aspect of kinetic proofreading' is used, as suggested by Hopfield (Hopfield, 1974 and 1978). The proposed mechanism of prebiotic translation can utilize ideale the time delay mechanism for proofreading. Firstly, there would be a mean time for which amino acid adenylate would have to wait for the formation of aminoacyl bond. This time would clearly be larger in non-enzymatic reactions and only the 'correct' amino acid would be able to stay for such a long time at anticodon; the 'wrong' one would leave off owing to weaker interaction. So the correct amino acid would have more chance of getting acylated. An amino acid once acylated and fixed on its position would wait for the next amino acid to come and occupy the neighbouring anticodon. This 'waiting period' would also be used as time delay in kinetic proofreading. If a wrong amino acid is acylated there is a chance that it falls apart before it finds a neighbour to join with because of weaker protection offered, against hydrolysis by anticodon. In this way by two step 'time delay kinetic proofreading' quite a good amount of fidelity would be achieved in prebiotic translation. As soon as a polypeptide of a critical size is formed, it would peel itself off the template and make it available for the next round of synthesis (thermodynamics of the process discussed below). In this manner, the proposed 'static template surface' would be able to direct translation in the absence of ribosomes and synthetases in prebiotic conditions.

3. 'Fossil' Evidences for Inter-Primal-tRNA Association

Evidences similar to the ones for Hopfield's structure are found to exist in favour of the hypothetical interactions also. The 5 bases long c regions, situated 8 bases away from the central base of the anticodon towards the 3′ end, of every tRNA were compared with e regions, consisting of 5 bases, 8 bases away from the central base of the anticodon towards the 5′ end (Figure 1(a)) for complementarity in *E. coli* (36 tRNAs). The same operation was performed with *B. subtillis* tRNAs (10 in number) also (sequences from Gauss and Sprinzl, 1981). Extra lengths in tRNAs having extra big 'lump' were looped out. It was found that on an average $\bar{n} = 1.85$ bases in *E. coli* and 2.34 in *B. subtillis*, out of 5 are complementry. If there had not been any reason for pairing, and if c and e regions of different tRNAs were independent of each other, only a $\bar{n} = 5(^1/_4)$ would be expected by just coincidence. To test whether the observed difference is significant, a chi-square analysis was performed, assuming a normal distribution for the random pairing with a mean of 1.25. The number of comparisons (36×36 in case of *E. coli* and 10×10 in case of *B. subtillis*) were high enough to assign a normal distribution to it. The null hypothesis that the observed frequencies are same as the one expected from random pairing could be rejected by 99.5% confidence from the chi-square values, in case of *E. coli* as well as *B. subtillis*.

There may be many ways in which $\bar{n}$ based on random assignments can exceed 1.25, for example if the entire tRNA contained G and C only. To test such a bias the frame of e was shifted in first case by one base toward 5′ end and in second case by two bases (now considering a sequence of 5 bases, 9 and 10 bases away from the central base of

the anticodon, respectively). In both the cases, the $\bar{n}$ observed was near 1.25, i.e., 0.96 and 1.28 respectively in case of *E. coli* and 0.76 and 1.33 in *B. subtillis*, consistent with the random pairing. Pairing of c and d of the same tRNA showed an average of $\bar{n} = 1.00$ in *E. coli*.

Different tRNAs have evolved quite free from each other and there is no correspondence between region c and e of different tRNAs. Then why do they still show such a pairing? This can be explained by the proposed hypothesis that primal-tRNAs might have been associated with each other through c and e regions, which were possibly fully complementary to each other before ribosomes and synthetases evolved. Despite the constant mutation and evolution, some residual ('fossil') pairing continues to exist in contemporary tRNAs.

The hypothesis implies that as one ascends in phyla the $\bar{n}$ should decrease. This was indeed found to occur in case of yeast, where $\bar{n}$ was 1.46 (the number of tRNAs analysed was 20). It was not possible to perform same analysis with higher organisms because sufficient number of sequences are not available for any single species.

The approximate date of existence of the proposed mechanism can be calculated if we know the rate of base substitution in tRNA. The rate of base substitution (w) is known to be roughly 1.6×10^{-10} replacements per site per year for several ribosomal RNA species (Hori *et al.*, 1977). Assuming that this rate is applicable to tRNA also (Hopfield, 1978), and that 5 out of 5 bases in c and e regions were paired at that time, the time (T) since appearance of ribosomes and synthetases can be given as (Kimura and Ohta, 1973):

$$T = -\frac{ln\,(1 - {}^{4}/_{3}\,\lambda)}{1.33\,w},$$

where λ stands for fraction of bases unchanged (1.85/5 in this case). T works out to be 3.18×10^9 yr, which would be slightly less in reality because the rate of base substitution would have been substantially higher in the early phase of life. This is consistent with the age of blue green algae, the organisms at the lowest evolutionary strata, which have been known to have existed for 2.5×10^9 yr (Schopf, 1970).

4. The Hypothesis is Compatible Thermodynamically

Base pairing between region c and e of different primal-tRNAs would make them interact cooperatively with the messenger molecule. The binding of one primal-tRNA molecule to a codon would facilitate binding of the next one at the adjacent site. The adaptors (tRNAs) would tend to cluster on the messenger and if the amount of primal-tRNAs is not sufficient to cover all of the messengers, some of them would be entirely covered but the rest would remain bare. this condition of *all-or-none limit* type binding can result only if τ, the equilibrium constant of inter primal-tRNA interaction is sufficiently high (of the order of 10^3). τ is given as (Bloomfield *et al.*, 1974):

$$\tau = \exp(-\frac{\triangle G}{RT})$$

where, $\triangle G$, is the free energy of interaction between adjacent ligands, minimum possible magnitude of which is -5 Kcal/mole, if all of the 5 bases are paired (Tinoco *et al.*, 1973). If the temperature is assumed to be 300 K, τ comes out to be 4023.9, which is sufficiently high. If slightly fewer bases are paired or there is some minor steric hindrance in the interactions, τ would be lesser, but is within the permissible range of *all-or-none* limit type cooperative binding. Thus, thermodynamically, the existence of static template surfaces is quite feasible.

Apparently it seems that the newly formed polypeptide would bind even more strongly to the 'static template surface' than the adaptors themselves do with the messenger molecule, because the peptide bond is stronger than the inter-primal-tRNA link, and the polypeptide will not be able to 'peel itself off' from the template. But this would not be so, because the extremely weak interactions (of the order of 0.5 Kcal/mole) between amino acid and anticodon can be readily interrupted by the random thermal motions. Before the formation of the peptide bond amino acid is kept on its place mainly by the aminoacyl linkage with the acceptor stem. The amino acid-anticodon interactions are only for the discrimination between 'wrong' and a cognate adaptor, which is achieved mainly through the 'time delay kinetic proofreading' as discussed before. When the peptide bond is formed the aminoacyl linkage is broken and resultantly any polypeptide is attached to the complex only through a single aminoacyl bond. As soon as the polypeptide size reaches a critical limit thermal movements of now bigger molecule become strong enough to break this bond too, and the polypeptide slips out of the complex.

5. Discussion

Although the details of the model refer to a comparatively later phase of the origin of life, it is continuous with its own origin which can be easily traced back. Small oligonucleotides (10–50 bases long) can faithfully replicate themselves without the help of any enzyme, just by complementary interactions (Eigen and Schuster, 1979; Lohrmann *et al.*, 1980). In the microspheres in which the hypothetical nucleotidyl polymerase activity was trapped (see Introduction) a good number of such oligomers is likely to have been present, showing low fidelity replication. This would result in a host of oligomers with limiting complementarity with each other. Such a population would organize itself into static template surfaces of a rudimentary nature, in which there is no strict periodicity of adaptors. This would evolve into the more refined structures, as the number of oligomers increases, because that would increase the availability of better suited adaptors in the system.

The model overcomes the two fundamental problems of ribosome-free prebiotic translation. Firstly, the free energy of codon-anticodon interaction (in water) alone is not sufficient to keep two adjacent adaptors attached sufficiently long enough for

peptide transfer to occur (Tinoco *et al.*, 1968). Secondly, there is the problem of 'back insertion'. This arises when a primal-tRNA with a nascent peptide in its 'lap', attached to a codon in the middle of the message is not able to decide whom to transfer its peptide to – to the next charged primal-tRNA on its right or on its left, because probabilities of both transfers are the same. This would result in a situation analogous to the phenomenon of 'random walk', in which the growing peptide, instead of forming linkage with the amino acid corresponding to the next codon accepts that of the preceeding one and the message is not read properly. The proposed hypothesis solves both of the problems without postulating improbable conformational changes and putting restrictions on the sequence of messenger (Crick *et al.*, 1976).The problem of back insertion does not arise here, and inter primal-tRNA association makes the binding with the messenger thermodynamically feasible.

The evolution from the proposed static template surfaces to ribosomes is also a continuous one. For the present purpose, the ribosome basically performs two functions in translation: (1) provides non-aqueous environment at the codons under consideration so that tRNAs can bind more tightly than is possible in aqueous phase (free energy of codon-anticodon interaction is far lower in non-aqueous surroundings than that of aqueous ones) (Tinoco *et al.*, 1968); (2) moves vectorially (5′–3′ direction) over the message, shifting its attention from one codon to the next, sequentially. In the evolution of ribosomes, first some protein aggregates might have been formed around the 'static template surfaces', providing them with a non-aqueous environment in which inter-primal-tRNA interactions were no longer required (because tighter codon-anticodon interactions were sufficient to keep the primal-tRNAs fixed to the message). Such protein aggregates later developed the capacity of energy dependent vectorial movement over the messenger molecule. In such a system, the primal-tRNAs left behind after the passage of the protein aggregate (now the rudimentary ribosome), would detach themselves owing to weaker interactions with codons in the new aqueous environment. Meanwhile, the new charged primal-tRNAs would enter the rudimentary ribosomes, attach themselves strongly to their respective codons (in the now non-aqueous environment) to contribute to the growth of the peptide chain. The evolution of charging of tRNA by aminoacyl synthetases possibly occurred simultaneously. Such a system would be able to evolve itself into the marvellous contemporary translational machinery.

If this hypothesis is assumed to be correct it implies that present genetic code was a three letter code even before the arrival of ribosome and synthetase (when the frame of e is shifted one or two bases away from the anticodon no residual pairing is observed, see earlier section), and there is some direct physical reason why the present three letter code was chosen.

The model removes one step of possible error in primitive translation, i.e., the attachment of the wrong adaptor to codon, because primal-tRNAs are already permanently attached to the messenger molecule. The other step where error can infiltrate into the system, i.e., binding of the wrong aminoacid to primal-tRNA is reduced considerably by '*mean waiting time*' (see earlier section) in which the 'wrong' amino acid is more open to hydrolytic attack.

In principle the hypothesis can be tested experimentally. Chemical synthesis coupled with cloning techniques can be used to prepare RNA molecules of the proposed secondary structure. Synthetic procedures, strong and fast enough to produce genes as long as that of insulin are available (Crea *et al.*, 1978). The exact physical conditions in which the static template surfaces would be formed and function can be predicted. It can be checked whether the complexes direct translation or not in such conditions.

Acknowledgements

I am very grateful to Prof. H. K. Das, Dr Prasanna Mohanty and Dr J. Subba Rao, for their sincere help in various stages of the work, and thankful to Indranil Dasgupta. Sreenivas Kumar, C. P. Geevan and Girija Ramakrishnan for their help in preparation of the manuscript. I wish to acknowledge the financial assistance provided by Rameshwardasji Birla Smarak Kosh in the form of a fellowship.

References

Balasubramanian, R., Seetharamulu, P., and Raghunathan, G.: 1980, *Origins of Life* **10,** 15.
Bernal, J. D.: 1951, *Physical Basis of Life*, Roultedge and Paul, London.
Bloomfield, V. A., Crothers, M. D., and Tinoco, I. Jr.: 1974, *Physical Chemistry of Nucleic Acids*, Harper & Row, New York.
Crea, R., Kraszewski, A., Hirose, T., and Itakura, K.: 1978, *Proc. Nat. Acad. Sci. USA* **75,** 5765.
Crick, F. H. C., Berner, S., Klug, A., and Pieczenik, K. G.: 1976, *Origins of Life* **7,** 389.
Eigen, M. and Schuster, P.: 1979, *The Hypercycle*, Springer-Verlag, Berlin.
Fox, S. W.: 1960, *Science* **132,** 200.
Fox, S. W. and Dose, K.: 1972, *Molecular Evolution and Origin of Life*, W. H. Freeman, San Francisco.
Gauss, D. H. and Sprinzl, M.: 1981, *Nucleic Acids Res.* **9,** rl.
Hiroyuki, O., and Egami, E.: 1979, *Origins of Life* **9,** 171.
Hopfield, J. J.: 1974, *Proc. Natl. Acad. Sci. USA* **71,** 4135.
Hopfield, J. J.: 1978, *Proc. Natl. Acad. Sci. USA* **75,** 4334.
Hori, H., Higo, K., and Osawa, S.: 1977, *J. Mol. Evol.* **9,** 191.
Jungck, J. R.: 1978, *J. Mol. Evol.* **11,** 211.
Kimura, M. and Ohta, I. T.: 1973, *Nature New Biol.* **243,** 199.
Kornberg, A.: 1974, *DNA Synthesis*, W. H. Freeman, San Francisco.
Lacey, J. C., Stephens, D. P., and Fox, S. W.: 1979, *Bio Systems* **11,** 9.
Liebl, V., Novak, V., Bejsovcova, L., Masinovskij, Z., and Oparin, A. I.: 1978, in H. Noda (ed.), *Origin of Life*, Japan Scientific Societies Press, Tokyo.
Lohrmann, R., Bridson, P. K., and Orgel, L. E.: 1980, *Science* **208,** 1464.
Miller, S. L. and Orgel, L. E.: 1974, *Origins of Life on Earth*, Prentice-Hall, London.
Paecht-Horowitz, M.: 1974, *Origins of Life* **5,** 173.
Paecht-Horowitz, M.: 1976, *Origins of Life* **7,** 369.
Rao, M., Odom, D. G., and Oró, J.: 1980, *J. Mol. Evol.* **15,** 317.
Sagan, C. and Ponnamperuma, C.: 1973, *Nature* **222,** 199.
Schopf, J. W.: 1970, *Biol. Rev.* **45,** 319.
Steinman, G. D., Lemmon, R. M., and Calvin, M.: 1965, *Science* **147,** 1574.
Tinoco, I., Jr., Davic, R. C., and Jaskunas, R. S.: 1968, in B. Pullman (ed.), *Molecular Associations in Biology*, Academic Press, New York.
Tinoco, I., Jr., Borrer, P. N., Dangler, B., Levine, M. D., Ohlenbeck, O. C., Crothers, D. M., and Rolloa, J.: 1973, *Nature New Biol.***246,** 40–41.
Weber, A. L. and Lacey, J. C., Jr.: 1978, *J. Mol. Evol.* **11,** 199.