# POLYNUCLEOTIDE REPLICATION COUPLED TO PROTEIN SYNTHESIS: A POSSIBLE MECHANISM FOR THE ORIGIN OF LIFE

A. G. MACKINLAY

*School of Biochemistry, University of New South Wales, Kensington, N.S.W. Australia*

(Received 5 May, 1981; in revised form 4 January, 1982)

**Abstract.** A mechanism is suggested for the replication under primitive conditions of long polynucleotides by the sequential incorporation of sequences related to those of modern transfer RNAs. It is proposed that replication of such molecules became established as the result of a replicative advantage arising from the concomitant linkage together of amino acids to form polypeptides. Initially these polypeptides may have been of random sequence. Selection of primitive tRNAs in which the amino acid and anticodon stem sequences were rotationally symmetrical could have led to specific, anticodon-directed aminoacylation and fixation of the genetic code along the lines suggested by Hopfield. (Hopfield, 1978). The primitive replication-coupled system would then have been able to synthesize specific proteins containing one amino acid residue for each primitive tRNA incorporated during replication. The end result of this line of evolution is postulated to have been a nucleoprotein structure resembling the ribosome. The primitive system would then have been able to give rise directly to triplet-coded protein synthesis. Some recent RNA sequence data are discussed which are consistent with derivation of modern protein synthesis from the primitive replication-coupled mechanism.

## 1. Introduction

Theories concerning the chemical events which led to the appearance on the primitive Earth of living systems need to explain how nucleic acid replication and protein synthesis began in the absence of the specific enzymes and nucleoprotein assemblies required for these processes today. Nucleic acid replication under these conditions is not too difficult to visualise due to the ability of complementary base sequences to spontaneously form base-paired structures under appropriate conditions. The problem of the origin of protein synthesis under primitive conditions is much more difficult. The essence of the problem is: how could a polynucleotide under prebiotic conditions specify with sufficient accuracy the assembly of a protein, e.g. a replicase, capable of facilitating the replication of that polynucleotide.

In a recent extensive analysis of the problem of the origin of life this relationship between polynucleotide and a polynucleotide-coded protein which acts back on the former to facilitate its replication, was referred to as hypercyclic coupling (Eigen and Schuster, 1977, 1978a, and b). These authors concluded that if a primitive hypercycle could once be established then evolution into complex organisms then becomes possible. The problem remains: to understand, on the basis of our knowledge of present-day organisms, how prebiotic molecules managed to establish the first primitive hypercycle.

In modern protein synthesis specificity depends on two steps: one is the aminoacyl-tRNA synthetase step in which each amino acid is linked to a specific tRNA as a

*Origins of Life* **12** (1982) 55–69. 0302–1688/82/0121–0055$02.25

result of recognition by an enzyme of both components of the reaction. The other step at which specificity is imposed is the codon-anticodon interaction between a triplet of bases in the mRNA and a complementary triplet in the tRNA anticodon loop. The equivalent interaction under primitive conditions poses a particular problem since interactions involving only three base pairs are not sufficiently stable in the absence of other stabilising elements of the modern translation machinery. It was pointed out by Crick (1968) that it would not be possible to maintain the specificity of protein synthesis in a situation where the coding ratio was changing. This has generally been taken as a strong argument in favour of protein synthesis having been triplet-coded from the outset. More recently it has been proposed that a mRNA-tRNA interaction involving five base pairs might have been compatible with decoding of a comma-less triplet code, provided there were restrictions on the nucleotide sequence of the primitive mRNA (Crick *et al.*, 1976; Eigen and Schuster, 1978b). This possibility notwithstanding, it remains very difficult to visualise how triplet-coded protein synthesis could have begun under strictly prebiotic conditions.

In this paper I present a speculative scheme wherein the first functional proteins were made by a mechanism in which the linkage of amino acids was part of the process of polynucleotide replication. The primitive mechanism depended on specific linkage of amino acids and tRNAs but did not involve a decoding step analogous to the codon-anticodon interaction. It is proposed that primitive replication-coupled protein synthesis led to the formation of a nucleoprotein structure resembling in essential respects the ribosome. At this stage triplet-coded protein synthesis became possible as a result of the appearance of tRNAs having conformations similar to those of modern tRNAs. Both mechanisms of protein synthesis then operated side-by-side until expression of the superior potential of triplet-code protein synthesis resulted in the uncoupling of replication from protein synthesis and the eventual displacement of the primitive protein synthesis mechanism.

## 2. Coupled Polypeptide and Polynucleotide Synthesis

In this section it is shown how, in principle, the coupling of polypeptide synthesis to the synthesis of a polynucleotide could result in a replicative advantage for the polynucleotide. Later it is shown how this effect could have functioned in a particular situation to generate an evolving replication/translation system.

Let it be supposed that under prebiotic conditions activated nucleotides were formed which were able to condense together to form oligonucleotides, which in turn condensed to form longer single-stranded polynucleotides. The formation of double-stranded polynucleotides might then have occurred as the result of the annealing of short oligonucleotides of complementary base sequence to the longer template strand. Condensation of these would give rise to an intact double-stranded structure. In principle such double-stranded structures could be replicated by separation of the intact strands followed by the annealing and condensation of oligonucleotides, as before, to yield two daughter double-stranded molecules.

If now the additional assumption is made that under conditions prevailing on the primitive Earth it was common for the 3′ end of the annealing oligonucleotides to be blocked by reaction with amino acids, then the replication mechanism outlined above would have been inhibited since condensation of the oligonucleotides would have been prevented. This inhibition could, however, have been relieved if a mechanism developed whereby transfer occurred of the blocking aminoacyl group of one oligonucleotide onto the aminoacyl group attached to an adjacent oligonucleotide. Repetition of such a process would allow the sequential condensation of oligonucleotides annealed along the template strand and at the same time would result in the formation of a polypeptide containing one amino acid for each oligonucleotide incorporated. Figure 1 illustrates this system of coupled polypeptide and polynucleotide synthesis. Such a system could be thought of as a hypercycle but in this case physical coupling rather than functional coupling occurs between polynucleotide and polypeptide synthesis.

The important conclusion to be stressed is that under particular conditions, namely those where 3′ aminoacylation of oligonucleotides occurs, a replicative advantage would accrue to the polynucleotide able to achieve peptide transfer in the fashion illustrated in Figure 1. This replicative advantage would depend not on the amino acid sequence of the polypeptide nor on its having any function or specificity, but upon peptide bond formation *per se*. If coupling between replication and non-specific protein synthesis were to cause replication of the polynucleotide concerned to become established there would then be the possibility that at a later stage the synthesis of functional proteins could begin. A particular situation where this may have been possible is described below.

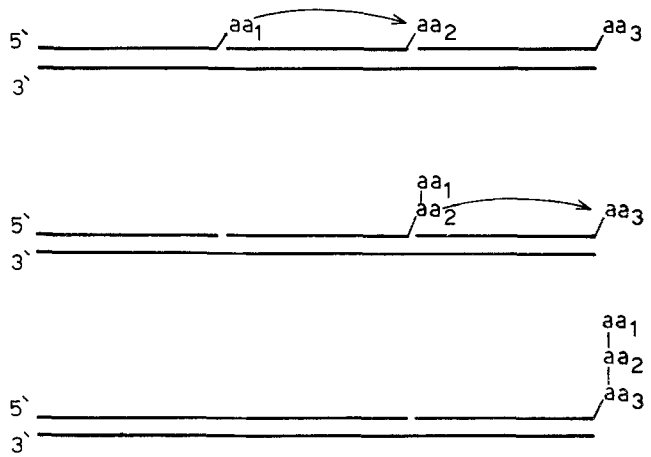


Fig. 1. Replication-coupled protein synthesis.

Under conditions where short oligonucleotides become aminoacylated at their 3′ termini, it is proposed that the conversion of a single-stranded polynucleotide into a double-stranded structure would be facilitated by concomitant protein synthesis. Transfer of the aminoacyl residue blocking the 3′ end of the first oligonucleotide onto the adjacent aminoacyl residue would allow condensation of the two oligonucleotides. Repetition of this process would yield a continuous polynucleotide strand with a polypeptide attached to its 3′ terminus containing one amino acid for each oligonucleotide incorporated.

## 3. Replication of Polynucleotides Containing Rotationally Symmetrical Sequences, with Coupling to Protein Synthesis

The concept of replication-coupled protein synthesis expounded above suggests that replication schemes involving tRNA-like molecules should be explored, since the aminoacylation of these polynucleotides is a central step in modern protein synthesis. The similarity of tRNAs in all present day organisms indicates that the tRNA structure must have been standardised at a very early stage during evolution and likewise suggests that the replication possibilities of this general type of molecule be focussed upon in considering primitive replication mechanisms. Arguments for the antiquity of tRNA structure may be found in Eigen and Schuster (1977 and 1978b).

In this section, therefore, I present a speculative mechanism for the protein synthesis-coupled replication under primitive conditions of long polynucleotides composed of repeating units of shorter polynucleotides similar to modern tRNAs. The general justification for this scheme is that, as will be detailed later, it appears to have the potential first of all to begin synthesizing small proteins of defined amino acid sequence and then to give rise directly to modern triplet-coded protein synthesis.

I begin by assuming that prebiotic reactions generated small polynucleotides similar to tRNAs in that both halves of the molecules possessed internal complementarities. Furthermore it is proposed that both halves of the primitive tRNAs (referred to here as ptRNAs) were themselves completely complementary. These ptRNAs could therefore adopt either the familiar cloverleaf conformation or exist as a hairpin loop i.e. they could undergo a lineform-cruciform transition (see Figure 2(a)). If one now constructs a long polynucleotide by condensing successive ptRNAs (Figure 2(c)) and converts this to a double stranced structure, the result is a molecule composed of successive palindromic sequences (Figure 2(d)).

Figure 3(a) illustrates the initial step in the suggested replication scheme of this double-stranded poly-ptRNA molecule. Separation of strands at one end of the molecule allows annealing of a ptRNA molecule to both arms of the resulting fork (Figure 3(a)). Subsequent migration of the fork allows annealing of the ptRNA molecule along the parental (−) strand with an equivalent displacement of the parental (+) strand (Figure 3(b)). Repetition of this process would result in the sequential incorporation of ptRNA molecules and result in semi-conservative, asymmetric replication of the original parental molecule. If the assumption is now made that the 3′ termini of the ptRNA molecules can be aminoacylated and furthermore that the aminoacyl residue of the first ptRNA may be transferred to the aminoacyl residue at the 3′ terminus of the second ptRNA to form a dipeptide (Figure 3(c)–(d)), then it follows that a polypeptide would be synthesized containing one amino acid for each ptRNA incorporated.

These assumptions are the same as were made before (Figure 1) to illustrate the general concept of protein synthesis coupled to polynucleotide replication. In the present case where aminoacylated ptRNAs are being incorporated, the formation of a peptide linkage between successive amino acids is seen to be feasible since both

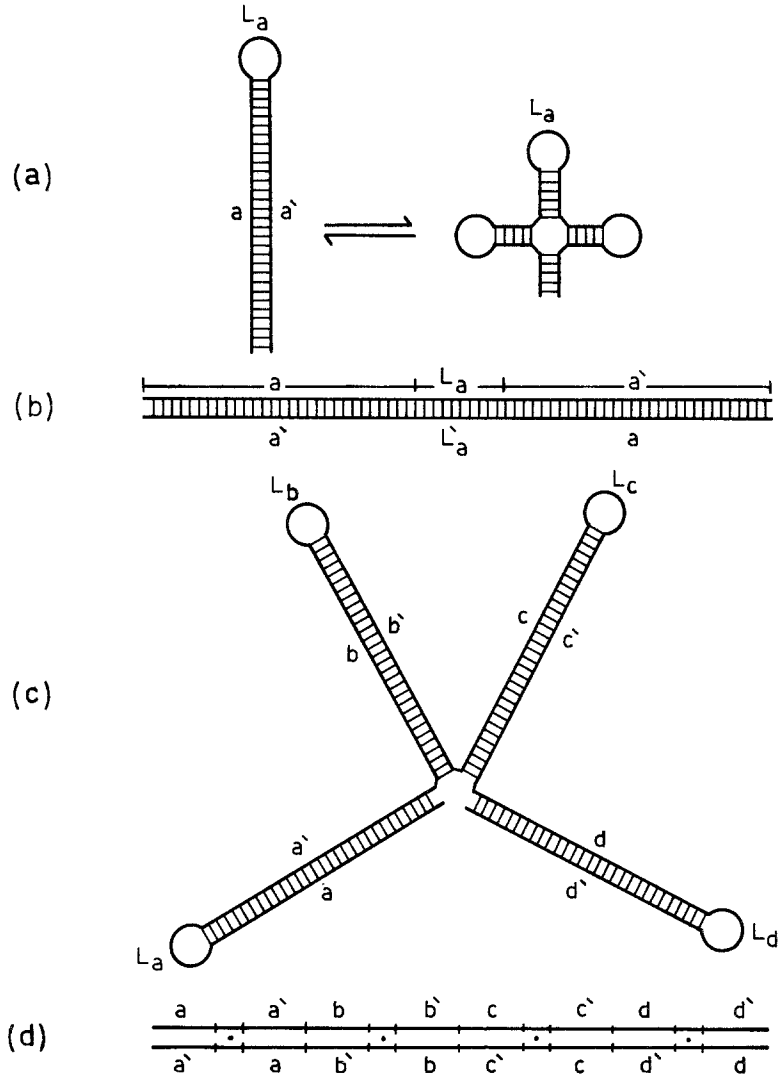


Fig. 2. Structure of ptRNAs and poly-ptRNAs.

(a) Lineform and cruciform configuration of ptRNA. The structure shown, containing about 65 nucleotides represents the minimum size which would be compatible with the formation of a cloverleaf structure with stable loops and stems. Otherwise there would be no restrictions on the number of bases in the loops and stems provided the 5′ and 3′ halves of the molecule retained their internal complementarities and their complementarity with each other. It is assumed, however, that most ptRNAs were roughly similar in size to modern tRNAs.

(b) Conversion of a ptRNA molecule to a completely double stranded structure would yield a molecule exhibiting hyphenated rotational symmetry. The centre of symmetry would be located at the midpoint of the $L_a$ loop sequence.

(c) Condensation of four different ptRNA molecules to yield a continuous single stranded polynucleotide referred to in the text as poly-ptRNA.

(d) Conversion of the poly-ptRNA shown in (c) to a completely double stranded structure which consists of an array of rotationally symmetrical sequences interrupted by non-symmetrical sequences corresponding to the single-stranded loops of the original ptRNAs.

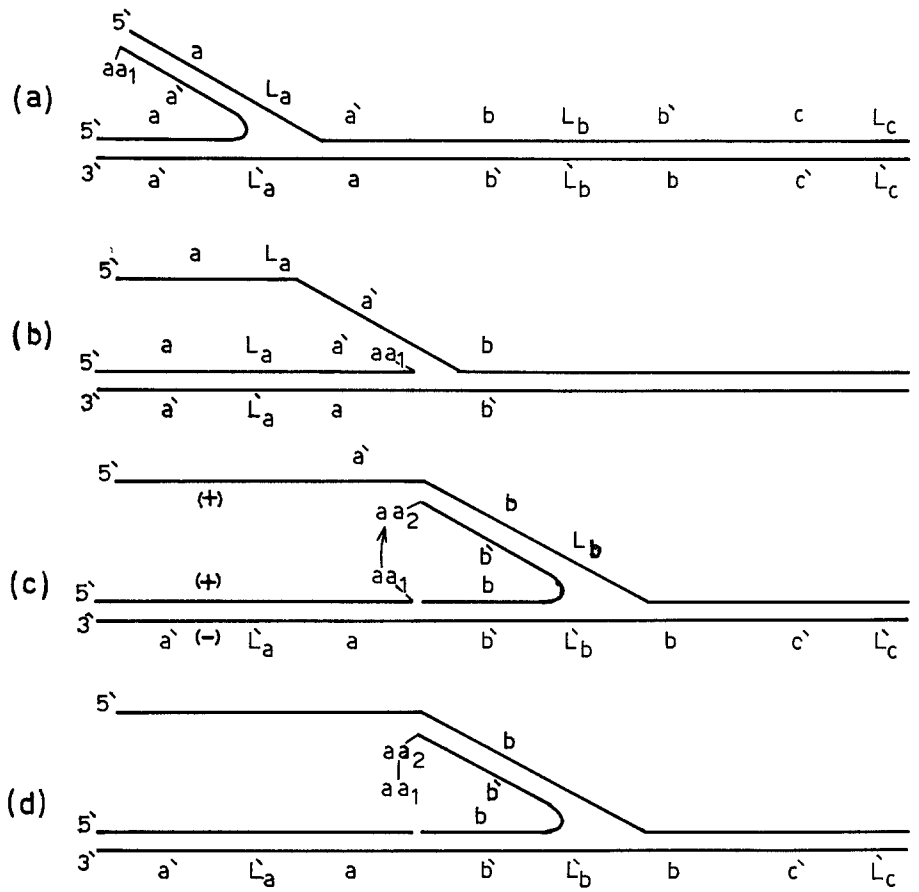


Fig. 3. Proposed mechanism of replication of double-stranded poly-ptRNA molecules.
(a) Separation of the strands at one end of the double-stranded poly-ptRNA allows the incorporation of an aminoacyl-ptRNA molecule containing the sequence a-$L_a$-a′ such that it base pairs with both the (+) and (−) parental strands to form a forked structure.
(b) Further separation of the parental strands allows annealing of the newly incorporated ptRNA along its full length to the parental (−) strand.
(c) Further displacement of the parental (+) strand allows another aminoacyl-ptRNA molecule to interact with the parental molecule to form a forked structure. It is proposed that the two newly incorporated aminoacyl ptRNAs are now so positioned that transfer of the first amino acid onto the second amino acid is now possible.
(d) The ptRNA occupying the fork now contains a dipeptide at its 3′ terminus. The process now repeats itself until the parental (+) strand is completely displaced. The daughter (+) strand generated by condensation together of successive ptRNAs will then bear a polypeptide at its 3′ terminus which will contain one amino acid for each ptRNA molecule incorporated.

amino acids are now situated close to each other. As before, peptide bond formation provides a replicative advantage by clearing the 3′ termini of incorporated ptRNAs and allowing them to condense into a continuous daughter (+) strand. An additional replicative advantage provided by coupling to protein synthesis would also apply to polynucleotides replicating by this particular mechanism due to the fact that the 3′

termini of ptRNAs occupying the replicating fork will at all times be blocked either by an aminoacyl or peptidyl group. This would serve to prevent growth of the 3′ end of the ptRNA strand along the displaced (+) strand which, should it occur, would result in a permanently forked structure unable to complete replication. A further replicative advantage enjoyed by molecules constructed from rotationally symmetrical rather than random sequences is that replication is accomplished by unzippering the parental molecule in stages from one end. For replication of a random sequence as in Figure 1 it was necessary to postulate complete strand separation followed by annealing of short complementary sequences.

An aspect of the replication scheme proposed here which has to be considered concerns the source of the ptRNAs which become incorporated during the asymmetric replication process. It has to be assumed that, at very early times, the source of ptRNAs was independent of the coupled replication process itself and that the ptRNAs required for replication were selected from a pool of similar structures which arose by purely prebiotic processes. Once the protein synthesis-coupled process had become established it would very rapidly have become important for the system to generate its own supply of ptRNAs in order to conserve advantageous sequences.

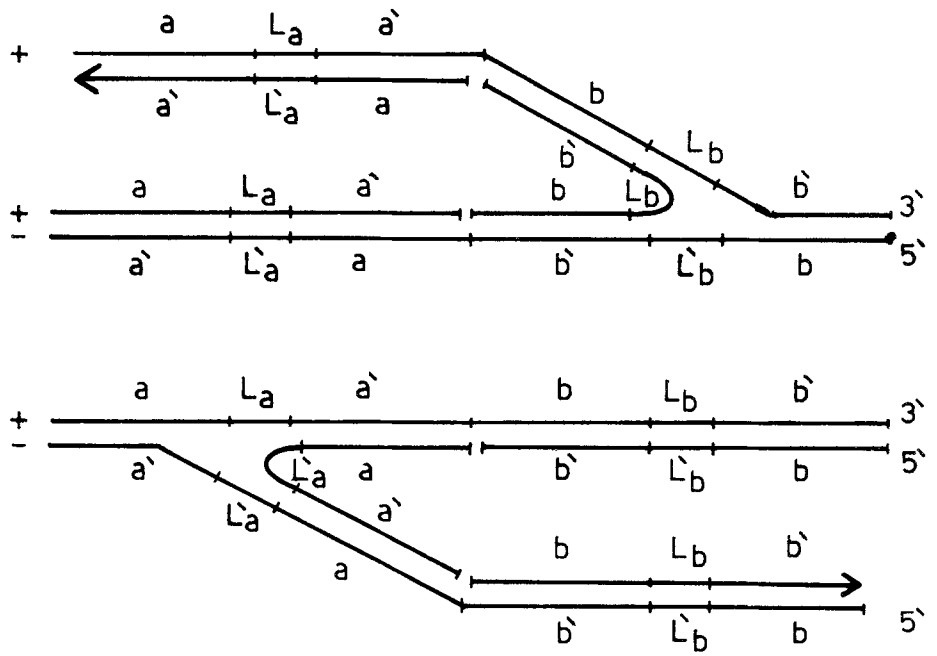


Fig. 4. Synthesis of new ptRNA molecules by back synthesis along the displaced parental strand. In the upper diagram a ptRNA of sequence a-$L_a$-a′ is synthesized using the parental (+) strand as template. This ptRNA is able to be incorporated during replication in the opposite direction, as shown in the lower diagram. The same is true for the ptRNA of sequence b-$L_b$-b′ which is is shown being synthesized in the lower diagram and being incorporated in the upper diagram. It is assumed that at early times the synthesis of ptRNAs occurred as the result of annealing and condensation of short oligonucleotides along the parental strand template.

This could have come about by using the displaced parental strand as template for the synthesis of new copies of ptRNAs (see Figure 4). It is necessary to postulate that copies synthesized in this way then dissociated from the parental (+) strand in order to be available for subsequent rounds of assymetric replication. One would expect dissociation to be facilitated by the lineform-cruciform transition which both the parental and copy strands would be able to undergo on account of their internal sequence complementarities.

If, as suggested, new ptRNAs were synthesized by annealing and condensation of oligonucleotides along the displaced (+) strand, it is apparent that the sequence a-$L_a$-a′ on the parental (+) strand would specify a ptRNA molecule of sequence a-$L'_a$-a′ (where the prime denotes complementarity). This ptRNA would be able to be incorporated during a subsequent round of replication starting from the opposite end of the molecule in which the (−) strand was displaced. Figure 4 illustrates how ptRNAs generated using one strand as template would be suitable for incorporation during displacement of the opposite strand. If replication occurs from both ends of the molecule, it would follow then that first the (+) strand and then the (−) strand would be displaced. These might then reanneal to regenerate the original parental double-stranded molecule. This would constitute a fully conservative replication mechanism.

## 4. The Assembly of Specific Amino Acid Sequences by The Primitive Replication-Coupled System

In the system described so far, replication is dependent on polypeptide synthesis but does not rely on these polypeptides having defined sequences or specific functions. In order for the system to be capable of further evolution it needs to be able to synthesize specific proteins. This requires that there be specificity in the aminoacylation step. Recent work by others suggests how this may have come about. Weber and Lacey (1978) and Jungck (1978) have shown that a relationship exists between the properties of the amino acids and those of their anticodon nucleotides. This indicates that in primitive systems aminoacylation specificity must have been determined at least in part by the nucleotide sequence of the anticodon. Hopfield (1978) performed a statistical analysis on sequences of modern tRNAs and found evidence for the existence at earlier times of complementarity between the 3′ side of the D stem and the sequence adjacent to the 3′ terminus of the molecule. Base pairing between these sequences would place the anticodon sequence in close physical proximity to the aminoacylation site. This hypothesis can be accommodated to the scheme presented here by simply requiring that from the population of replicating molecules selection occurred for those molecules able to undergo anticodon-directed aminoacylation. These were ptRNAs in which the sequences corresponding to the amino acid and anticodon stems were rotationally symmetrical. Figure 5 illustrates the sequence symmetries of such a molecule and how these would permit the ptRNA to assume a dumb-bell configuration equivalent to that proposed by Hopfield. Provided aminoacylation

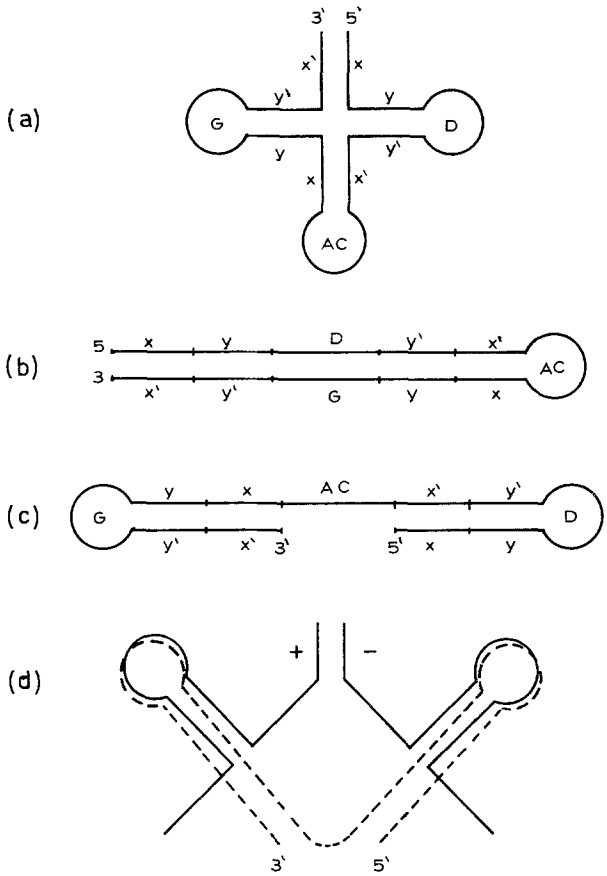


Fig. 5. Sequence and symmetry characteristics of ptRNAs capable of anticodon-directed aminoacylation.
(a) A ptRNA molecule drawn in the conventional cloverleaf configuration. The sequences x and x′ in the amino acid and anticodon stems are rotationally symmetrical. The sequences in the D and G stems are similarly related.
(b) The ptRNA molecule shown in (a) drawn in the hairpin loop configuration (cf Figure 2(a)).
(c) The ptRNA molecule shown in (a) drawn in the dumb-bell configuration in which the 3′ terminus is located in the vicinity of the sequence corresponding to the anticodon triplet of modern tRNAs. The ability of the anticodon stem sequence, x, to base pair with the amino acid stem sequence, x′, was postulated by Hopfield (1978) to permit anticodon-directed specific aminoacylation of primitive tRNAs.
(d) Suggested interaction between the (+) and (−) strands of the parental molecule and an incoming ptRNA molecule.
Interaction between the D and G loops of the ptRNA with complementary unpaired loop sequences of the parental molecule would possibly stabilise the ptRNA in the dumb-bell configuration required for specific aminoacylation. The complex would then transform to the structure shown in Figure 3.

occurred while the ptRNA was in this form, its specificity would be determined by the anticodon sequence and nearby surrounding nucleotides and fixation of the genetic code would result. Therefore in order to account for aminoacylation specificity in the primitive system proposed here, I rely on Hopfield's hypothesis and the reader should consult his paper for further details.

In order to minimise the chance of hydrolytic removal of the aminocyl group, aminoacylation is postulated to occur at the replication fork immediately prior to incorporation of the ptRNA. Figure 5(d) illustrates a possible initial complex formed between the ptRNA in the dumb-bell configuration and the arms of the parental molecule. The latter are shown looped out to form single strand loops complementary to the *G* and *D* loops of the incoming ptRNA. Base pairing between these unpaired regions of the parental molecule and of the ptRNA might be expected to stabilise the latter in the dumb-bell configuration required for aminoacylation. This would be followed by annealing of the ptRNA and parental strands to form the replication complex shown in Figure 3.

The primitive system of replication-coupled protein synthesis would therefore have been able to make proteins of defined sequence. The sequences made would have been dictated by the order in which ptRNAs were linked together in poly-ptRNAs. Within a large population of poly-ptRNAs some would give rise to proteins able to facilitate the various steps involved in replication-coupled protein synthesis and so promote replication of the poly-ptRNA which specified them. The primitive system with a 'coding ratio' of 60–70 to 1 would have made only small proteins perhaps containing 50–100 amino acids, if the size of present-day replicating RNA molecules is any guide. These proteins would have been large enough to evolve catalytic activity even if, as enzymes, they were inefficient by modern standards. They would certainly have been large enough to fulfil structural roles similar to those of the modern ribosomal proteins.

It has often been suggested that the early steps in the appearance of life took place at a clay or mineral surface or interface. Some such surface would appear to have been necessary for the establishment of the system of replication-coupled protein synthesis suggested here. Perhaps an apatite mineral, able to discriminate between single and double stranded structures, may have provided a suitable surface for a replicating complex in which the arms of the replicating fork undergo lineform-cruciform transitions. Whatever the nature of this inorganic component, it is apparent that there would have been a great evolutionary advantage available if replication-coupled protein synthesis was able to evolve in such a way as to become independent of the inorganic component. I propose that this occurred when a nucleoprotein particle arose which was similar in essential respects to the modern ribosome. The RNA component of this particle would have corresponded to one or more single stranded poly-ptRNA molecules generated by displacement during replication. In fact this RNA would be in the direct line of descent of modern ribosomal RNAs. The protein components would have arisen as products of the primitive protein synthesis system and would have evolved functions which facilitated operation of the primitive, coupled system. Their role in the primitive nucleoprotein particle (which I will refer to here as the protosome, in order to distinguish it from the ribosome) would have been similar to the role of modern ribosomal proteins.

The proposal is that the protosome served to provide a surface on which replication-coupled protein synthesis was able to proceed, in the same general way as the

ribosome contributes to modern protein synthesis. The consequences of this development was that this primitive form of life was now no longer dependent upon and confined to mineral surfaces, but would have been able to colonise other environments.

## 5. The Transition to Triplet-Coded Protein Synthesis

In order for the aminoacylated ptRNA to become incorporated at the replication fork it must anneal to the parental strands as shown in Figure 3. The initial interaction, it was suggested above, was between the ptRNA in dumb-bell configuration and the looped out parental strands (Figure 5(d)). Following insertion of the ptRNA at the replication fork, looping out might occur as shown in Figure 6(a). Looping out of the replication complex in this fashion may have been necessary to bring the two amino acids sufficiently close together for peptide bond formation to occur. The point which is emphasized in Figure 6(a) is that, in order to accommodate the looped out segments of polynucleotide, there would have evolved, on the surface of the protosome, binding sites for the unpaired sequences generated by looping out. Figure 6(a) depicts such binding sites, one occupied by the G loop of the ptRNA, the other by the complementary unpaired sequence from the displaced parental strand. Figure 6(b) shows that these binding sites could also serve to allow binding of conventional tRNAs (in their modern conformation) in positions allowing codon-anticodon interactions. That is, the binding site on the small subunit of the protosome/ribosome would correspond to the binding site for peptidyl-tRNA or, during initiation, for the initiating methionyl-$tRNA_f^{Met}$. The binding site on the large subunit would correspond to the amino acid site of the modern ribosome.

The critical development which allowed triplet-coded protein synthesis to begin may have been the appearance of a polymerase activity able to carry out replication of ptRNAs. This polymerase, a product of the replication-coupled protein synthesis system could have evolved in response to the requirement to provide a source of ptRNAs for incorporation into poly-ptRNAs. (This requirement was initially satisfied by annealing of oligonucleotides to the displaced parental strand – see Figure 4). The availability of this polymerase would have led to the propagation of variant ptRNAs, having conformations resembling modern tRNAs. These, as a result of their variant sequences would have been unable to be incorporated into poly-ptRNAs by the primitive replication mechanism and would therefore have been available to initiate triplet-coded protein synthesis as shown in Figure 6(b). A further assumption which has to be made is that the specific aminoacylation which operated in the primitive system was carried over into the modern system. Thus a constraint upon the evolution of these variant ptRNAs would have been the requirement that they continue to be specifically recognised by primitive aminoacyl-tRNA synthetases.

Immediately following the advent of triplet-coded protein synthesis the situation would have been that both forms of protein synthesis proceeded together for an extended period. On the one hand, the primitive system would be producing relatively

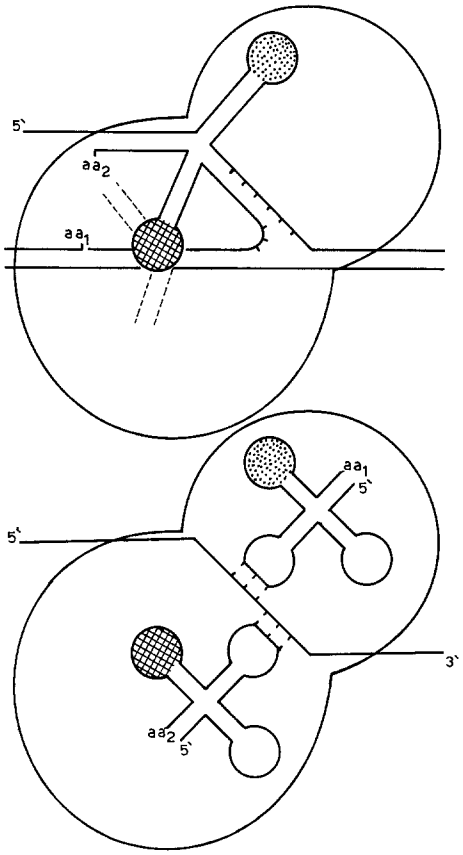


Fig. 6. Proposed relationship between the primitive protosome and the modern ribosome.
(a) Positioning of the aminoacyl-ptRNA at the fork is dependent upon both arms of the fork adopting a cruciform structure (cf Figure 5). The GTΨC loop (loop IV) of the ptRNA is depicted interacting with a binding site on the large subunit of the protosome. The complementary single strand loop generated by the displaced parental (+) strand binds to a similar site on the small subunit.
(b) The binding of two aminoacyl-tRNAs to the subunits of the ribosome during triplet-coded protein synthesis depends on their interaction with the same binding sites illustrated in (a). Positioned in this way the anticodons of both tRNAs base pair with codons on the mRNA. The latter traces a path across the ribosome similar to that followed by the displaced (+) strand on the protosome. It is assumed that the three-dimensional structure of the tRNAs is such as to position the amino acids at sites similar to those occupied by the amino acids on the protosome but no attempt has been made to indicate this on the diagram. For both figures it is necessary to imagine the small subunit folded over on top of the large subunit to create a cleft or interface.

simple proteins but proteins which had undergone extensive evolution leading to acquisition of a variety of functions, such as the polymerase activity, aminoacyl-tRNA synthetases, and a variety of protein components of the protosome. On the other hand, triplet-coded protein synthesis would have been able to produce larger proteins but initially these proteins would have had little or no function. Due to their larger size their potential for the acquisition of function would have been much greater and it would have been inevitable that eventually they would displace proteins

made by the primitive system. During the transition period there is no obvious reason why both mechanisms of protein synthesis would not have been entirely compatible.

As the products of triplet-coded protein synthesis evolved the function of the protosome/ribosome would have reduced to protein synthesis alone. This would have been accompanied by extensive modification of its RNA components as well as replacement of the protein components with modern ribosomal proteins. Nevertheless those structural features involved in protein synthesis, namely the tRNA binding sites and the secondary structure of rRNA essential for ribosome function, would be expected to have been strongly conserved since their maintenance would have been a continuing essential requirement. Aminoacyl-tRNA synthetases produced by the primitive protein synthesis system would have been replaced by triplet-coded synthetases, however the genetic code would have remained fixed as before.

## 6. Discussion

The foregoing account of the origins of protein synthesis involves the operation at early times of a mechanism of protein synthesis which was subsequently supplanted and for which direct evidence is unlikely to be obtainable. There are however, a number of correspondences between mechanisms and structures postulated here, and characteristics at the molecular level of present-day living systems. Some of these are briefly surveyed below.

A variety of evidence points to interrelationships between RNA replication and protein synthesis. Engelberg and Schoulaker (1976) compared nucleotide sequences from MS2 and Q$\beta$ RNAs with portions of *E. coli* 16S rRNA and found extensive homologies. Two of the components of Q$\beta$ replicase are the host-specified tRNA-binding proteins Tu and Ts (Blumenthal *et al.*, 1972) while a third subunit of this enzyme is the ribosomal protein S1. (Wahba *et al.*, 1974). The ability of tRNA to serve as primer in the replication of some RNA viruses (Harada *et al.*, 1975; Rich and Raj Bhandary, 1976) may also be interpreted as reflecting an evolutionary relationship between RNA replication and protein synthesis.

The scheme proposed here states that rRNAs evolved from molecules which were in turn derived from molecules (ptRNAs) which were also the ancestors of modern tRNAs. This could account for the fact that rRNA sequences can be represented as a series of hairpin loops (Ehresmann *et al.*, 1975) and for the interspersion of tRNA genes with ribosomal genes (Lund and Dahlberg (1975); Anderson *et al.*, (1981). The evolutionary split between the rRNAs from eubacteria, archaebacteria and eukaryotes (Woese and Fox, 1977) can be accounted for on the basis that these three groups of rRNAs evolved from separate poly-ptRNAs. DNA sequence determinations show that many non-translated sequences, often involved in protein recognition, exhibit rotational symmetry. It seems an attractive explanation of their origin that these sequences have derived from primitive, rotationally symmetrical ptRNA sequences which, prior to the advent of the first polymerase, were the only sequences able to be replicated.

Although the emphasis in this paper has centred on the origins of protein synthesis, it is nevertheless apparent that the primitive replication mechanism could have given rise directly to modern polynucleotide replication mechanisms. The asymmetric displacement of parental strand (Figure 3) is a mechanism used by some modern viruses (Scott *et al.*, 1977). Back synthesis along the displaced (+) strand to provide ptRNAs for the primitive system (Figure 4) parallels the discontinuous back synthesis of Okazaki fragments during modern DNA replication.

There is one specific consequence of the proposals made here for which available data appear to provide support. This concerns the sequences which interact with the two binding sites on the protosome. According to the scheme derived here, these sequences were complementary (Figure 6(a)). Comparison of the base sequence in the GTΨC loop of initiating $tRNA_f^{Met}$, which binds at or near the peptide site, with that of chain elongating $tRNA_m^{Met}$, which binds at the amino acid site, indicates that this complementarity has been preserved.

In mouse myeloma cells the GTΨC loop sequences for $tRNA_f^{Met}$ and $tRNA_m^{Met}$ are:

$tRNA_f^{Met}$: 5′ GGAUCGAA 3′ (Piper and Clark, 1974;)
$tRNA_m^{Met}$: 3′ CCUAGCΨT 5′ (Piper, 1975)

These sequences are drawn antiparallel to indicate exact complementarity extending over an 8-base sequence. The corresponding sequences of *E. coli* $tRNA^{Met}$ also exhibit significant complementarity:

$tRNA_f^{Met}$: 5′ GGTΨCAAAUCC 3′

(Cory *et al.*, 1968)

$tRNA_m^{Met}$: 3′ CCUAAGCΨTGG 5′

None of the above considerations provide better than circumstantial evidence in favour of the theory proposed. The main hope for progress in understanding how life originated is that it will be possible to create experimental systems which display properties to be expected of primitive replication and protein synthesis systems. The scheme presented here indicates that attention should be directed at systems in which the two processes can be coupled and it makes fairly specific predictions about the kinds of polynucleotides likely to be involved. A major unknown in any such investigation concerns the nature of the inorganic surface postulated to have facilitated replication at very early times. The report that montmorillonite clays are able to polymerise alanine from alanyladenylate (Paecht-Horowitz *et al.*, 1970) suggests that it would be fruitful to investigate the ability of this material to interact with polynucleotides resembling the postulated ptRNAs and poly-ptRNAs and to promote replication-coupled polymerisation of amino acids.

## References

Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., and Young, I. G.: 1981, *Nature* **290**, 457.

Blumenthal, T., Landers, T. A., and Weber, K.: 1972, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 1313.

Cory, S., Marcker, K. A., Dube, S. K., and Clark, B. F. C.: 1968, *Nature* **220**, 1039.

Crick, F. H. C.: 1968, *J. Mol. Biol.* **38**, 367.

Crick, F. H. C., Brenner, S., Klug, A., and Pieczenik, G.: 1976, *Origins of Life* **7**, 389.

Ehresmann, C., Stiegler, P., Mackie, G. A., Zimmerman, R. A., Ebel, J. P., and Fellner, P.: 1975, *Nuc. Acid Res.* **2**, 265.

Eigen, M. and Schuster, P.: 1977, *Naturwiss.* **64**, 541.

Eigen, M. and Schuster, P.: 1978a, *Naturwiss.* **65**, 7.

Eigen, M. and Schuster, P.: 1978b, *Naturwiss.* **65**, 341.

Engelberg, H. and Schoulaker, R.: 1976, *J. Mol. Biol.* **106**, 709.

Harada, F., Sawyer, R. C., and Dahlberg, J. E.: 1975, *J. Biol. Chem.* **250**, 3487.

Hopfield, J. J.: 1978, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4334.

Jungck, J. R.: 1978, *J. Mol. Evol.* **11**, 211.

Lund, E. and Dahlberg, J. E.: 1977, *Cell* **11**, 247.

Paecht-Horowitz, M., Berger, J., and Katchalsky, A.: 1970, *Nature* **228**, 636.

Piper, P. W. and Clark, B. F. C.: 1974, *Eur. J. Biochem.* **45**, 589.

Piper, P. W.: 1975, *Eur. J. Biochem.* **51**, 283.

Rich, A. and Raj Bhandary, U. L.: 1976, *Ann. Rev. Biochem.* **45**, 805.

Scott, J. F., Eisenberg, S., Bertsch, L. L., and Kornberg, A.: 1977, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 193.

Wahba, A. J., Miller, M. J., Niveleau, A., Landers, T. A., Carmichael, G. G., Weber, K., Hawley, D. A., and Slobin, L. I.: 1974, *J. Biol. Chem.* **249**, 3314.

Weber, A. L. and Lacey, J. C.: 1978, *J. Mol. Evol.* **11**, 199.

Woese, C. R. and Fox, G. E.: 1977, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5088.