# Killing two birds with one stone: a chemically plausible scheme for linked nucleic acid replication and coded peptide synthesis

## John D Sutherland[1] and Jonathan M Blackburn[2]

**To understand how life began, we must explain the origins of nucleic acid replication and genetically coded peptide synthesis. Neither of these is easy to explain individually; here, we propose a chemically plausible scheme for the evolution of a process that simultaneously produced both polymers. Later, two separate machineries could have evolved from the linked process.**

Addresses: [1]The Dyson Perrins Laboratory, South Parks Road, Oxford OX1 3QY, UK. [2]Department of Biochemistry, Tennis Court Road, Cambridge CB2 1QW, UK.

Correspondence: John D Sutherland
E-mail: john.sutherland@dpl.ox.ac.uk

It is widely accepted that evolution began with a prolonged phase in which RNA acted as a purveyor of genetic information, a replicase and a catalyst of a simple metabolism (Figure 1). It is proposed that the molecules comprising this 'RNA world' [1,2] then acquired the ability to produce genetically coded peptides and, later, DNA. For this to have happened, there must have been a robust way to make reasonably long oligonucleotides and then a selectable means of producing genetically coded peptides must have evolved. The first of these requirements is demanding and has stimulated many theoretical solutions and a few experimental investigations. But before dealing with the consequences of these experimental investigations, let us consider the second requirement.

One might imagine that genetically coded peptide synthesis would have arisen by a mechanism like that shown in Figure 2. But in considering this mechanism, one is immediately confronted by an apparent paradox [3]. The complex apparatus required for coded peptide synthesis has many parts, none of which is apparently useful individually, and yet the system must have evolved in a series of selectable steps. It is possible that RNA could fulfil the functions of the ribosomal proteins and the aminoacyl-tRNA synthetases. The synthesis of functionally competent proteins requires a reasonable set of aminoacyl-tRNAs,

## Figure 1
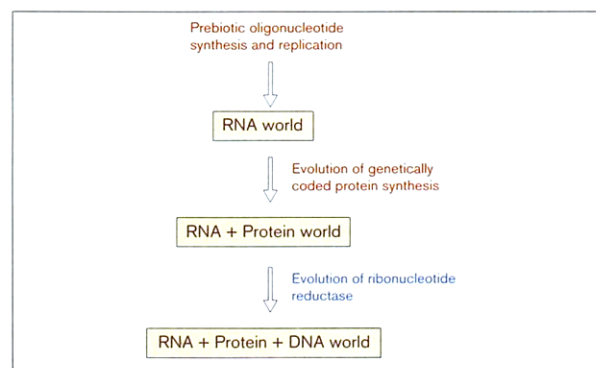


The current view of early evolution (which we challenge in this article) envisions the evolution of RNA, coded peptides and DNA in three distinct phases. The earliest phase, the 'RNA world', is supported by the fact that RNA can be both genetic material and catalyst [1,2]. But there are no plausible models for the evolution of coded peptide synthesis in isolation, and so to reach the second phase we are forced to postulate RNA-directed aminoacyl-tRNA synthesis and the chance evolution of a primitive ribosomal apparatus composed of RNA. Deoxyribonucleotides are currently synthesised by enzymatic reduction of ribonucleotides, which is taken to indicate that DNA arose after RNA and coded peptides; the DNA/RNA protein world is therefore shown as the third phase.

## Figure 2
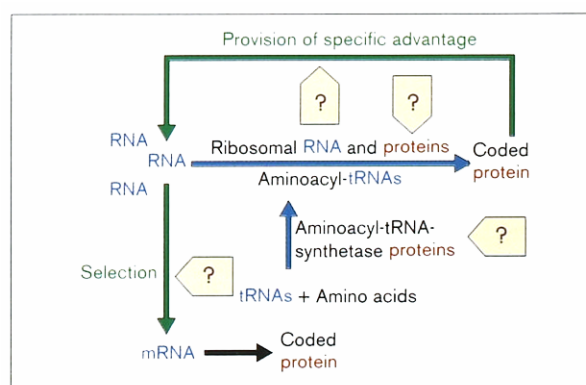


The evolution of genetically encoded peptide synthesis in an RNA world. This scheme appears paradoxical: protein synthesis requires proteins. To resolve this paradox, some have invoked specific RNA-catalysed aminoacylation of tRNAs. But it is hard to derive an advantage from protein synthesis unless a reasonably full set of specifically aminoacylated tRNAs is available. Specifically aminoacylated tRNAs might arise for another purpose and be subverted to protein synthesis, but no such plausible purpose has been suggested. Furthermore, if many RNAs are present, how does an encoded peptide specifically enhance the replication of the RNA that encoded it?

however, which could be acquired only if each member of the set conferred an advantage on the evolving RNA organism. This advantage could not have been the production of homopolymeric peptides (or peptides composed of only a few different amino acids), because such peptides are unlikely to have any specific function. Furthermore, to produce such a peptide the evolving RNA would have to contain long stretches of specific codons. This would impose an additional, rigorous constraint on the nature of prebiotic nucleic acid synthesis.

One apparent solution is to suggest that specific RNA aminoacylation could have been progressively selected for a different purpose and that, when a sufficient set of aminoacyl-tRNAs was available, a change of function allowed coded peptide synthesis [4]. This is an attractive approach to the problem, but leads to yet another difficulty: what was the original purpose for the aminoacyl-tRNAs? It is well-nigh impossible to imagine any function other than protein synthesis that might coincidentally require cognate amino acid attachment to the acceptor stem some 70 Å distant from the specific anticodon.

With no plausible theory for the evolution of coded protein synthesis available for testing, researchers have been unable to bring experimental chemistry to bear upon this crucial subject. As mentioned above, however, a number of theories concerning the prebiosis of nucleic acids have been addressed experimentally. The purpose of this article is to suggest that the by-products of one proposed prebiotic route to nucleic acids could be used for interdependent coded protein synthesis and nucleic acid replication and to show how this system could evolve divergently to the contemporary situation.

## Potential prebiotic routes to nucleic acids

Figure 3 shows three contrasting possibilities for the prebiotic synthesis of the natural nucleic acids: nucleoside triphosphate polymerisation, isomerisation of pyranosyl-RNA and aldol polymerisation of acyclic phosphodiesters. Modern nucleic acid replication involving enzymatic polymerisation of preformed nucleoside triphosphates (Figure 3a) is supremely efficient and is the ultimate target for any evolutionary scheme. Attracted by this, Orgel and others (e.g. [5,6]) have performed an extensive investigation of the nonenzymatic polymerisation of activated nucleotide derivatives. Although these elegant and important experiments have successfully demonstrated polymerisation, there are many problems with this approach. These include strong sequence dependencies, the requirement for very specific activation, the phenomenon of enantiomeric cross-inhibition, competing monomer hydrolysis, 2'/3' linkage ambiguity, product inhibition and the need to explain where the activated nucleotides came from (see [7] for review). Ingenious *in vitro* evolution experiments (e.g. [8]) have shown that ribozymes are capable of catalysing the formation of a

phosphodiester bond between short oligonucleotides, but not, as yet, a bond between two mononucleoside triphosphates. The fact that nature now uses a system in which mononucleosides are added to a growing nucleotide chain cannot necessarily be taken to mean that a noncatalysed version of this system was its evolutionary forerunner. It does mean, however, that any alternative scheme for the prebiogenesis of the natural nucleic acids must embody a reasonable mechanism for transition to this system.

Pyranosyl-RNA (Figure 3b,c) has been studied by Eschenmoser and coworkers ([9] and references therein) as part of an investigation of the 'chemical etiology of nucleic acid structure'. As an isomer of RNA in which base-pairing more strictly follows the Watson–Crick mode, pyranosyl-RNA has many fascinating properties *per se* and especially in the context of prebiotic chemistry. Pyranosyl-RNA tetramer-2,3-cyclophosphates can ligate nonenzymatically in a template-dependent manner without major sequence restrictions other than complementarity. The copying of pyranosyl-RNA templates is highly regioselective and has recently been shown to be chiroselective; homochiral templates mediate the ligation of tetramer units provided that the tetramer units are of the same chirality as the template. In addition, the ligation of homochiral tetramers is not inhibited by diastereoisomeric tetramers. The copolymerisation of a racemic mixture of tetramers containing all possible diastereoisomers is expected to give rise to stochastic libraries largely composed of individually homochiral oligomers. This work demonstrates that homochirality can arise spontaneously when a statistical ensemble of molecules polymerises to a stochastic ensemble of polymers (a point which has been touched-on elsewhere [10]). From the prebiotic perspective, the two remaining properties of pyranosyl-RNA that remain to be demonstrated are its capacity for self-assembly and its ability to undergo a chemically dramatic but information-retaining isomerisation to RNA. Certainly the transition (with retention of sequence information) is mechanistically plausible and a wonderful idea, but it would only explain how RNA arose. The conundrum of the evolution of the genetic code for peptide synthesis remains.

Our laboratory has studied the aldol polymerisation of acyclic phosphodiesters (Figure 3d) as a route to nucleic acid formation [7,10,11]. Oligomerisation of monomers based on the natural, aromatic bases to produce aldol dimers and trimers in the absence of template has been demonstrated (although the exact chemical nature of the oligomers remains to be established). Extensive experimental work remains to be done in this area but, so far, our results indicate that monomers containing dihydro forms of the bases are better candidates for natural nucleic acid production.

Although we know little that is certain about the prebiotic environment, we can at least be sure that the molecules

**Figure 3**
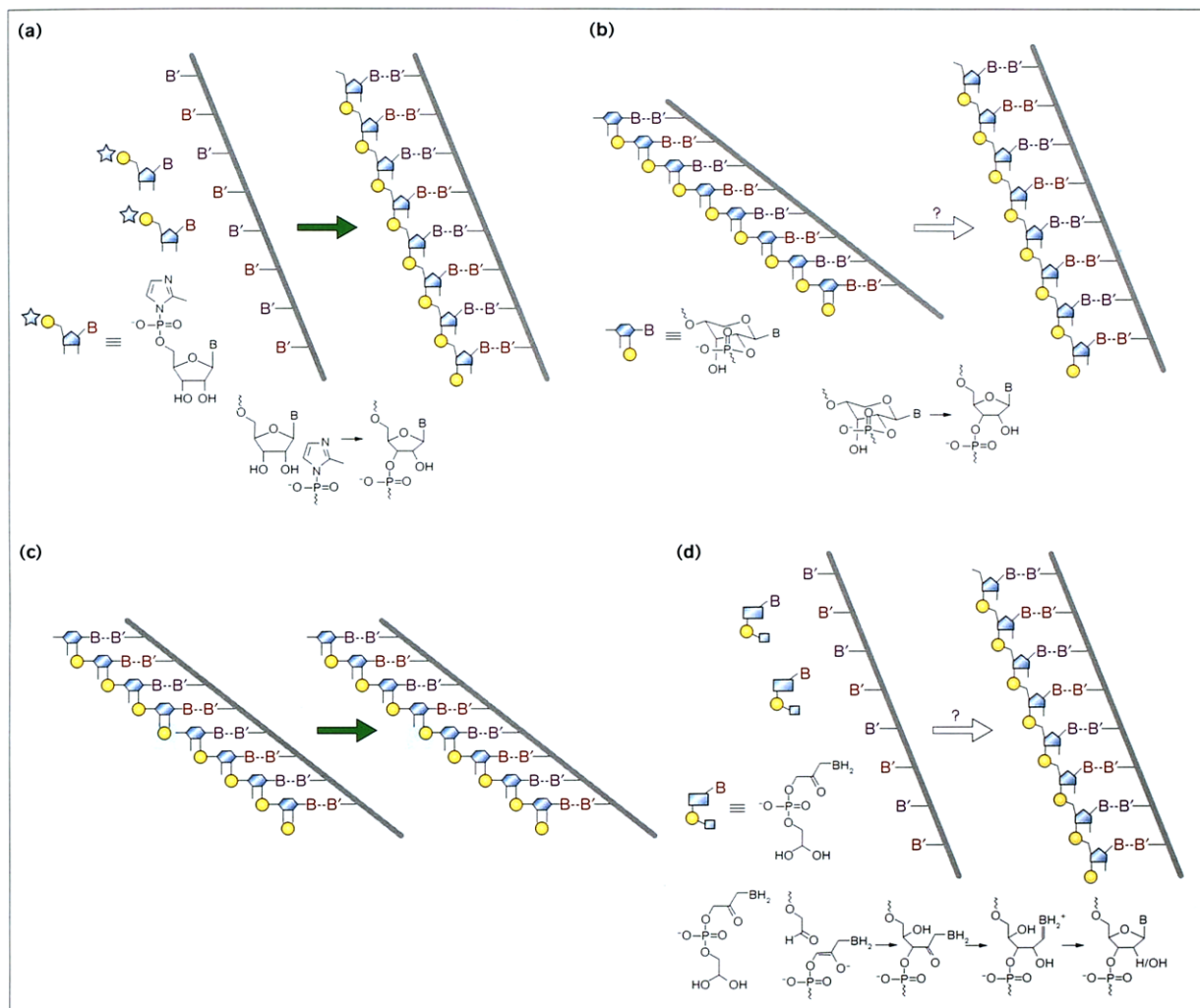


Three suggestions for the prebiotic synthesis of the natural nucleic acids. (a) Template-directed polymerisation of activated nucleotides. Although demonstrable *in vitro*, this is not considered to be a plausible prebiotic route to natural nucleic acids ([10]; see text for details). (b) The duplex isomerisation of pyranosyl-RNA to RNA. The properties of pyranosyl-RNA, for example its ability to show chiroselective template-directed ligation of tetramer-cyclophosphates [9] as shown in (c), are remarkable. If a plausible route for self-assembly can be found, this nucleic acid must be considered as a possible precursor to RNA, but its ability to isomerise to RNA remains to be shown. (d) The aldol polymerisation of acyclic phosphodiesters. Our investigations are much less advanced than those for (a) and (b)/(c), but early results using monomers containing aromatic bases are encouraging. Polymerisation in the absence of template is possible, but the required redox-transfer and ring-closure steps may be prevented by the restricted availability of the nitrogen lone pair in the attached base. A modification involving monomers containing dihydro-bases is shown; this might, in theory, give polymers containing both ribonucleotides and deoxyribonucleotides [7,11]. The possibility of linked RNA and DNA synthesis is thus suggested.

from which life evolved were components of a complex mixture. Applying this principle to the synthetic scheme shown in Figure 4 suggests that we should consider the effects of other compounds that might arise from the process used to produce the acyclic phosphodiesters, such as glycoaldehyde triphosphate and the amino acid derived dehydrooxazine-2-one [7]. Both of these compounds have the potential to act as chain-terminators for

the polymerisation of the dihydro-base monomers. It is thus possible that the polymerisation we invoke would not only result in long nucleic acid polymers but that short, apparent failure sequences capped at the 5′ end with a triphosphate and at the 3′ end with an aminoacyl group would also accumulate. Below, we propose that these aminoacyl-trinucleotide-triphosphates might allow both nucleic acid template replication and coded peptide synthesis.

**Figure 4**
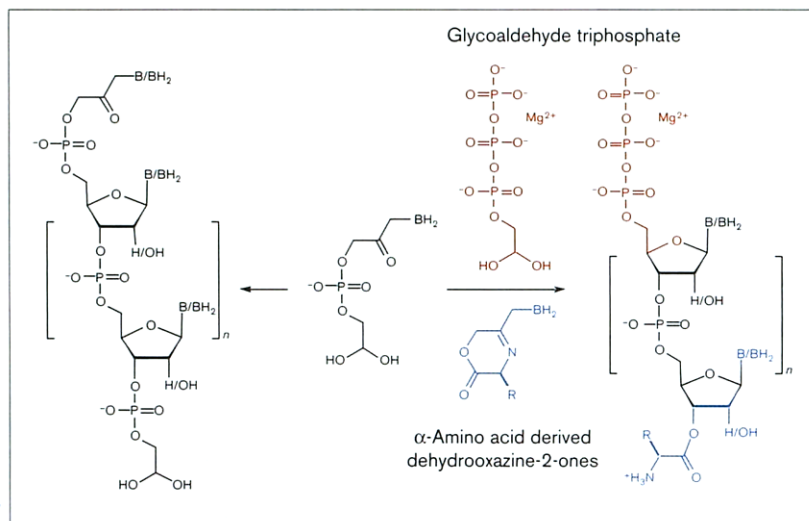


The relevance of 'contaminants' to the aldol polymerisation of acyclic phosphodiesters. Potential aldol donors and acceptors likely to be present in the prebiotic medium include glyco-aldehyde triphosphate (shown in red) and a variety of α-amino acid derived dehydrooxazine-2-ones (shown in blue), both of which are good candidates for intervention in the aldol polymeri-sation. Chemically, it is reasonable to suppose that the enamine tautomer of a dehydrooxazine-2-one would be a good aldol-type donor and the aldehyde group of glycoaldehyde triphosphate would be a good aldol acceptor. Polymerisation of a mixture containing these various compo-nents would therefore be expected to produce, in addition to polymeric nucleic acids, short terminated oligonucleotides. These terminated oligonucleotides would necessarily bear 5'-tri-phosphate and 3'-aminoacyl groups. Such chain termination is a disadvantage from the point of view of the robust prebiotic synthesis of long polymeric nucleic acids, but here we propose that the terminated by-products might have been crucial for the evolution of life.

## Linking fortuitous coded peptide synthesis and nucleic acid replication

Ribozymes can mediate oligonucleotide ligation, and pyranosyl-RNA tetramers can undergo nonenzymatic replicative oligomerisation. Taken together, these obser-vations suggest that template-directed ligation of short, suitably activated oligomeric nucleic acids is intrinsically easier than the polymerisation of similarly activated
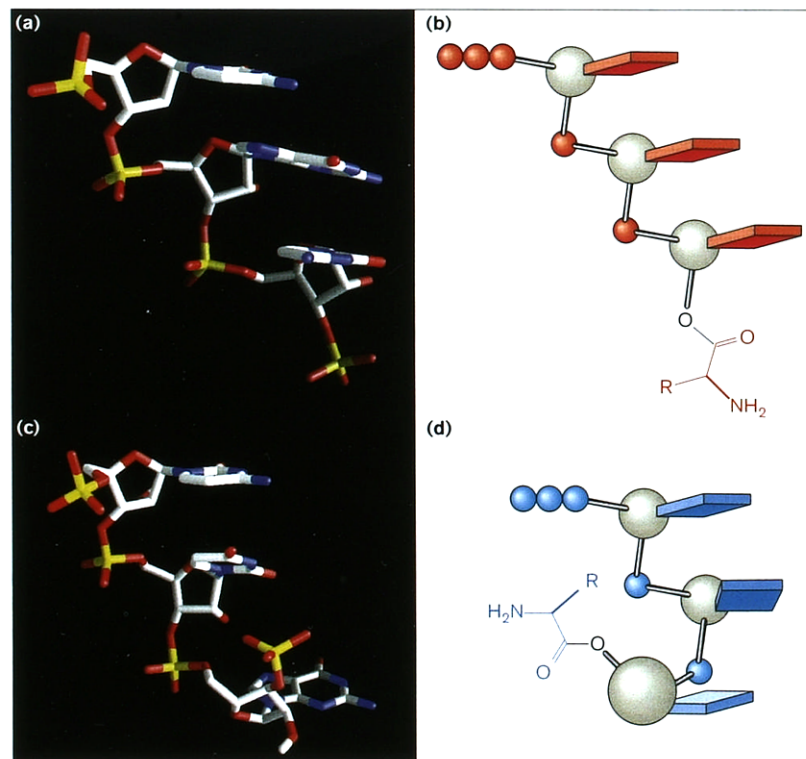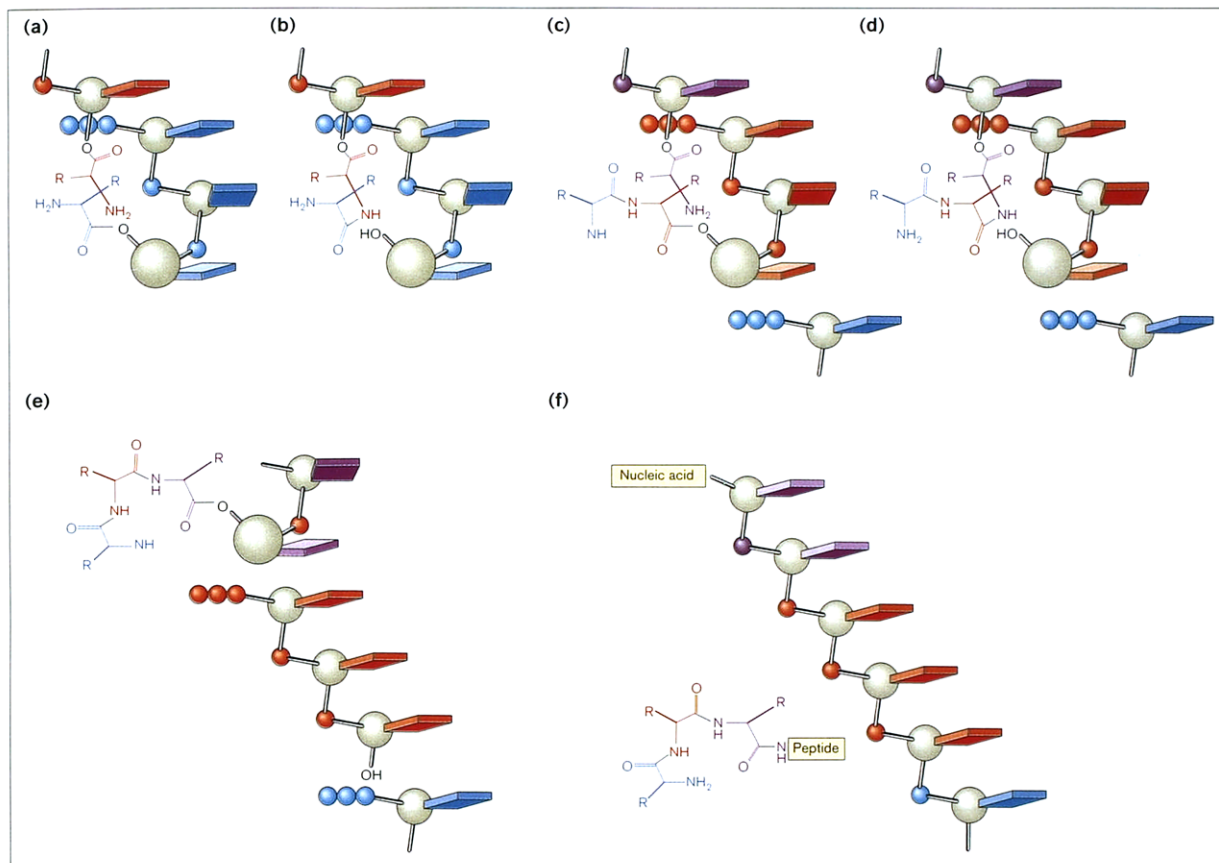
**Figure 5**



Trinucleotide conformational duality [12]. **(a)** The structure of an extended trinucleotide sub-fragment of tRNA<sup>Phe</sup> (residues 70–72); **(b)** a schematic representation of a 3'-aminoacyl-trinucleotide triphosphate in an extended conformation similar to that shown in (a). Two aminoacyl trinucleotides bound to a template in this extended conformation would clearly be unable to perform aminoacyl transfer to give one deprotected trinucleotide and one dipeptide trinucleotide. **(c)** The structure of the anticodon loop π-turn of tRNA<sup>Phe</sup> (residues 31–33); **(d)** a schematic representation of a 3'-aminoacyl-trinucleotide triphosphate in a folded-back conformation that is reminiscent of the conformation shown in (c) and is, therefore, clearly plausible. Model-building studies indicate that a trinucleotide in this conformation would be able to perform the aminoacyl transfer discussed above to a second trinucleotide in the extended conformation.

**Figure 6**



A scheme to link coded peptide synthesis and nucleic acid replication. **(a)** When a folded-back 3'-aminoacyl-trinucleotide triphosphate (blue) is brought by templating into contact with an extended one (red), a peptide bond can form **(b)**. **(c)** Fold-back of the dipeptidyl-trinucleotide triphosphate (red and blue) allows dipeptidyl transfer to a third 3'-aminoacyl-trinucleotide triphosphate (purple), forming a tripeptide **(d)**. **(e)** A conformational change allows the second uncharged trinucleotide triphosphate (red) to adopt an extended conformation in which its 3'-hydroxyl group is ideally positioned for ligation to the first uncharged trinucleotide triphosphate (blue). Ligation and iteration of the process produces a coded peptide and a copy of the nucleic acid template **(f)**. This scheme provides a plausible solution to many problems associated with separate coded peptide synthesis and nucleic acid replication; the phenomenon of 'wobble' is anticipated; amino acid sidechain contact with the dinucleotide base pair is possible only if the natural relative stereochemistry is followed; evolutionary advantage automatically accrues to the template encoding the peptide in that production of the peptide and replication of the template are intimately associated; the need for external ribosomal apparatus is avoided; and the requirement for enzymatic synthesis of aminoacyl-RNAs is avoided (they are in fact produced coincidentally as prebiotic aldol polymerisation by-products). Our proposal differs from previous suggestions concerning the origin of the genetic code ([13–16] and references therein) and is based on chemical reasoning coupled with the principle of evolutionary continuity [17].

monomers. We consider modified trimers here because of the favourable compromise between template-binding affinity and selectivity afforded by this oligomer length. In the case of the natural nucleic acids, it is easy to imagine that template-directed ligation of trinucleotide triphosphates could occur with a relatively simple peptide or RNA catalyst whereas the polymerisation of nucleoside triphosphates would be more demanding. For the ligation of trinucleotide triphosphates to occur, however, free 3'-hydroxyl groups are essential. In the scenario described above, any trinucleotide triphosphates would be 3'-aminoacylated and therefore unable to ligate unless somehow deprotected. If the deprotection of one template-bound aminoacyl-trinucleotide triphosphate were to be effected by the aminoacyl-amino group of an abutting aminoacyl-trinucleotide triphosphate then the first trinucleotide triphosphate would become competent for ligation and the second trinucleotide triphosphate would be converted to its 3'-dipeptidyl derivative. In essence, the ability to perform ligation could be a consequence of peptide-bond formation. An obvious requirement for the aforesaid peptide-bond formation would be that the nucleophilic amino group and

**Figure 7**
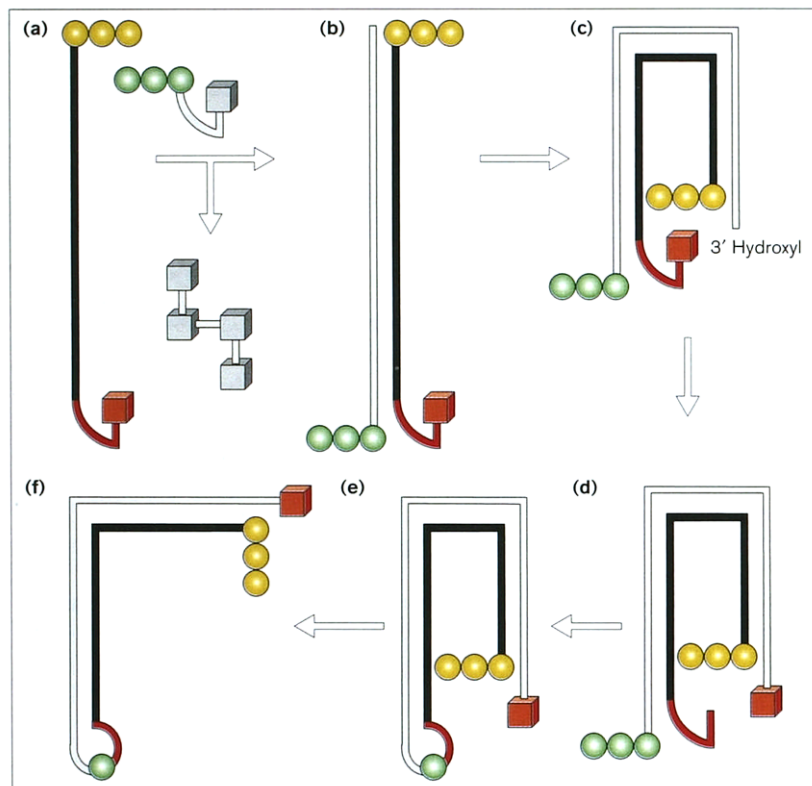


Proposed evolution of aminoacyl-tRNAs. Translation/replication of a template with a folded-back cognate 3′-aminoacyl terminus **(a)** results in a near complete duplex **(b)**. If the duplex can adopt a conformation in which the 3′-hydroxyl group of the newly synthesised strand can be aminoacylated by the 3′-aminoacyl group of the complementary strand **(c)** through transesterification, then the duplex shown in **(d)** becomes competent for ligation to form a hairpin **(e)**. After ligation, the resultant molecule will bear a 3′-aminoacyl group (red cube) which is cognate to the now spatially separated anticodon loop (red line). Such molecules will necessarily be 3′-amino-acylated hairpin-type structures with the anticodon situated approximately halfway through the sequence **(f)**, properties still found in modern aminoacyl-tRNAs.

the electrophilic aminoacyl-carbonyl partner should be close together. This is not possible when both oligonucleotides are in extended conformations, but model building shows that when the trinucleotide triphosphate aminoacyl-donor is in a folded-back conformation and the trinucleotide triphosphate aminoacyl-acceptor is in an extended conformation then the corresponding ester and amino groups are sufficiently close to allow a reaction. Examination of the nucleic acid structural data base revealed that the π-turn found in a variety of RNA molecules (including, tantalisingly, the anticodon loop of tRNA!) constitutes just this sort of folded-back conformation (Figure 5).

If an L-aminoacyl-group is modelled onto the 3′-hydroxyl group of a partly duplex π-turn trinucleotide conformation, then the aminoacyl sidechain is placed in contact with the major groove faces of the two 5′ base pairs. This contact would result in binding if there were any complementarity between the sidechain and the base pairs. It is difficult to match particular sidechains and dinucleotide base pairs but it seems unlikely that such an intimate association would not result in some specificity of binding. If a 3′-aminoacyl-trinucleotide triphosphate in an extended conformation is modelled to abut this folded-back 3′-aminoacyl-trinucleotide triphosphate, then it becomes

apparent that the geometry of the complex would be such as to allow aminoacyl transfer (Figure 6).

The fact that a conformation allowing aminoacyl transfer between trinucleotide triphosphates also forces contact between the aminoacyl sidechain and the dinucleotide base pairs of the folded-back conformer appears to us more than coincidental. If this contact confers specificity, then the binding of specific aminoacyl-trinucleotide triphosphates to a complementary template will be enhanced and any subsequent peptide-bond formation will result in coded aminoacyl transfer. Repetition of the process results in a peptidyl-trinucleotide triphosphate bound, by base pairing, to the 3′ end of the template and aligned trinucleotide triphosphates competent for replicative ligation. Hydrolysis of this terminal peptidyl-trinucleotide triphosphate would then produce free peptide. The production of the coded peptide confers a specific advantage to the encoding template in the sense that it makes replication of the template possible by allowing ligation of the aligned, complementary trinucleotide triphosphates (Figure 6). If one such coded peptide can itself act as a ligase, then this system will gain a very powerful selective advantage. At this evolutionary point coded peptide synthesis and nucleic acid replication are tightly linked. Continuation of the coupled

**Figure 8**

Proposed evolution of the amino acid code. **(a)** A schematic representation of the anticodon loop in the aminoacyl-tRNA formed as shown in Figure 7. The sequence derived from the folded-back 3'-aminoacyl-trinucleotide terminus of the template is shown in red; the first two 5' residues of the complementary strand are shown in purple. The aminoacyl group transferred to the acceptor stem was chemically 'cognate' to B1 and B2. If B1 and B2 are subsequently used in template recognition, then the new code will be the same as the old code. If the fold-back base, X, is used as a 'wobble' base and B1' and B2' are used for template recognition, then the new code will be a transposition of the old code. **(b)** The anticodon loop of tRNA<sup>Phe</sup> (residues 31–35). After the evolution of enzymatic aminoacylation of tRNAs there would be no requirement for B1' and B2' to be
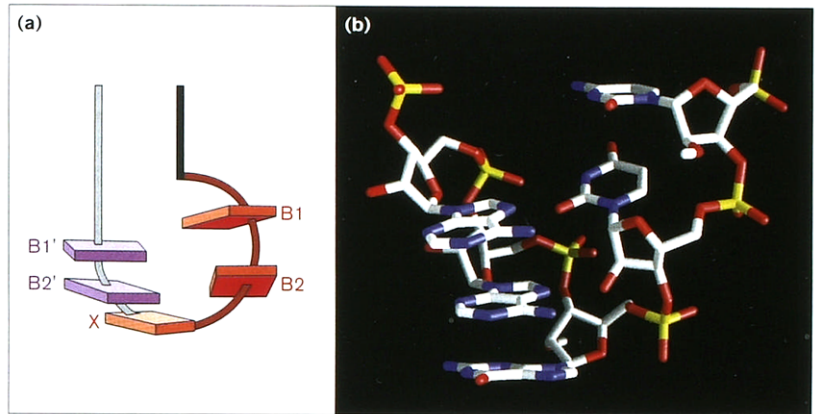


complementary to B1 and B2. The modern use of the fold-back base as the 'wobble' base at the 5'-terminus of the anticodon

suggests that a transposition of the genetic code is a real possibility.

process would eventually result in the pool of cognate aminoacyl-trinucleotide triphosphates becoming depleted; it would then become advantageous to shift to the separate machineries of peptide synthesis and nucleic acid replication found in contemporary biochemistry.

## Breaking the link

The proposed separation of protein synthesis from nucleic acid synthesis requires a way to produce a full set of specifically aminoacylated-tRNAs and a nucleoside triphosphate polymerase. We now propose a chemically plausible way this might have happened.

Aminoacyl-tRNAs might be produced via the scheme shown in Figure 7. Here, the unpaired portion of a template with a folded-back cognate aminoacyl-3'-terminus is copied, with concomitant peptide synthesis, to give a

structure that can readily rearrange into a looped conformation charged with an amino acid at the 3' end (Figure 7). A full set of aminoacyl-tRNAs might arise in this way. Note that the 3' end of the original template (which presumably confers some specificity for amino acid binding) is found in the centre of the final molecule, and would therefore naturally form part of the anticodon loop (see B1 and B2 in Figure 8). Residues complementary to the original template (B1' and B2' in Figure 8) are also present in the anticodon, and might have been used for template recognition, resulting in an inversion or transposition of the original code. Indeed there is some suggestive evidence (the existence of 'wobble bases') that this did indeed occur. The possibility of this transposition needs to be borne in mind in future attempts to establish molecular recognition between amino acid sidechains and dinucleotides.

**Figure 9**

Proposed evolution of a nucleoside triphosphate polymerase from a template-directed trinucleotide phosphate ligase. **(a)** We propose that trinucleotide ligation might be possible with a relatively simple (coded) peptide catalyst. The proximity effects required for ligation should, to a large part, be inherent to the templated substrates, and phospho-diester-bond formation might then require only general acid/base catalysis. Repeated rounds of catalysis would lead to complete replication of the template. **(b)** Minor modifications to the peptide might then make it possible to bring (otherwise weakly) templated mononucleoside triphosphates into proximity with the growing templated chain. This peptide, if capable of repeated rounds of catalysis, would be a primitive polymerase.
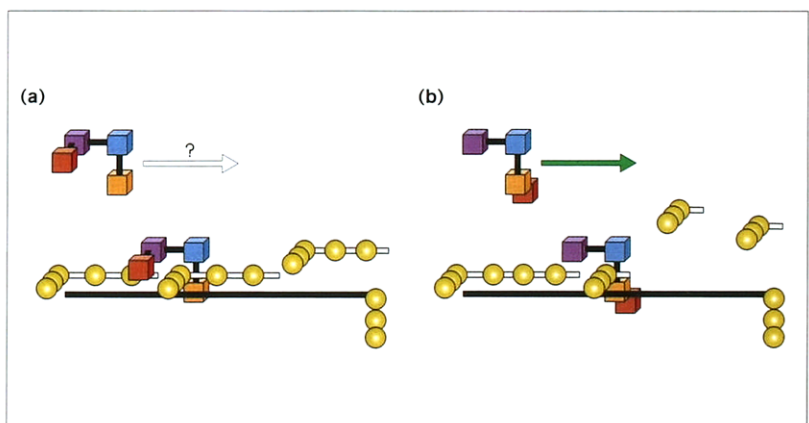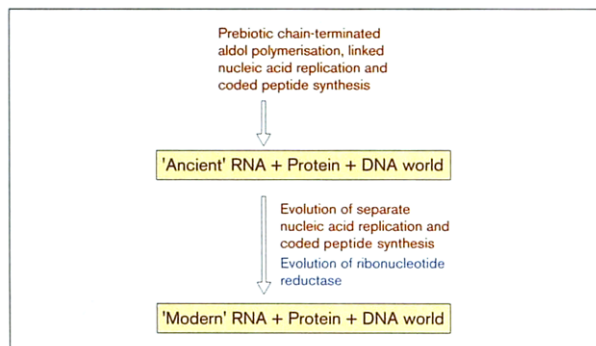
**Figure 10**



The scheme for early evolution suggested in this article. In contrast to the conventional view presented in Figure 1, this scheme entails the coordinated production of RNA, DNA and coded peptides. Divergent evolution then allows these systems to separate.

As these aminoacyl-tRNAs began to be used in an improved peptide synthesis, the evolution of a separate nucleic acid replication system would become possible, because replication would no longer be a direct consequence of peptide synthesis. In fact, evolution of either a separate method for peptide synthesis or a separate method for nucleic acid replication would drive the evolution of the other. The coded peptide capable of template-directed trinucleotide triphosphate ligation postulated above could evolve into a nucleoside triphosphate polymerase, without change of chemistry, by alterations in binding (Figure 9). In principle this seems likely to be an easy transition.

In summary, we suggest that early evolution might never have involved a purely RNA world (Figure 10). In our model, polypeptides coevolved with nucleic acids, and the 'fossil' RNAs that led to the proposal of the RNA world were simply RNA machines that functioned sufficiently well not to need replacement by a protein counterpart.

**Prospects**

We have suggested here that the severe problems associated with independent evolution of coded peptide synthesis and nucleic acid replication can be removed if the two processes are linked. This proposal provides chemically plausible explanations for the origin of coded peptide synthesis, the genetic code, the phenomenon of third base codon wobble and nucleic acid replication. Although initially linked, coded peptide synthesis and nucleic acid replication can be separated in a chemically plausible way, leading to a situation reminiscent of contemporary biochemistry. These ideas are speculative, but we air them because they are clearly capable of being put to test by experiment. We hope that fellow chemists and biologists will join in this endeavour.

**References**

1.  Gesteland, R.F., & Atkins, J.F., eds. (1993). The RNA World. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
2.  Gilbert, W. (1986). The RNA World. *Nature* **319**, 618.
3.  Orgel, L.E. (1968). Evolution of the genetic apparatus. *J. Mol. Biol.* **38**, 381-393.
4.  Weiner, A.M. & Maizels, N. (1987). t-RNA-like structures tag the 3' ends of genomic RNA molecules for replication: implications for the origin of protein synthesis. *Proc. Natl Acad. Sci. USA* **84**, 7383-7387.
5.  Ferris, J.P., Hill A.R., Jr., Liu, R. & Orgel, L.E. (1996). Synthesis of long prebiotic monomers on mineral surfaces. *Nature* **381**, 59-61.
6.  Inoue, T. & Orgel, L.E. (1983). A non-enzymatic RNA polymerase model. *Science* **219**, 859-862.
7.  Sutherland, J.D. & Whitfield, J.N. (1997). Prebiotic chemistry: a bioorganic perspective. *Tetrahedron*, in press.
8.  Wright, M.C. & Joyce, G.F. (1997). Continuous evolution of catalytic function. *Science* **276**, 614-617.
9.  Bolli, M., Micura, R. & Eschenmoser, A. (1997). Pyranosyl-RNA: chiroselective self-assembly of base sequences by ligative oligomerization of tetranucleotide-2',3'-cyclophosphates (with a commentary concerning the origin of biomolecular homochirality). *Chem. Biol.* **4**, 309-320.
10. Sutherland, J.D. & Weaver, G.W. (1994). Studies on a potentially prebiotic synthesis of RNA. *Tetrahedron Lett.* **35**, 9105-9108, footnote 21.
11. Sutherland, J.D. & Whitfield, J.N. (1997). Studies on a potentially prebiotic synthesis of RNA. *Tetrahedron*, in press.
12. Sussman, J.L. *et al.*, & Kim, S.-H. (1978). Crystal structure of yeast phenylalanine tRNA. I. Crystallographic refinement. *J. Mol. Biol.* **123**, 607-630.
13. Hopfield, J.J. (1978). Origin of the genetic code: a testable hypothesis based on tRNA structure, sequence and kinetic proof reading. *Proc. Natl Acad. Sci. USA* **75**, 4334-4338.
14. Shimizu, M. (1982). Molecular basis for the genetic code. *J. Mol. Evol.* **18**, 297-303.
15. Szathmary, E. (1993). Coding coenzyme handles: a hypothesis for the origin of the genetic code. *Proc. Natl Acad. Sci. USA* **90**, 9916-9920.
16. Rodin, S., Rodin, A. & Ohno, S. (1996). The presence of codon-anticodon pairs in the acceptor stem of tRNAs. *Proc. Natl Acad. Sci. USA* **93**, 4537-4542.
17. Crick, F.H.C. (1968). The origin of the genetic code. *J. Mol. Biol.* **38**, 367-379.