Artikel: Polarity at onset of genetic coding. I. Bipolar bondings in two-step takeover of peptide templates by prenucleic-ribonucleic acids

Autor: Turian, Gilbert

DOI: https://doi.org/10.5169/seals-740424

Archs Sci. Genève	Vol. 49	Fasc. 3	pp. 213-227	Décembre 1996

POLARITY AT ONSET OF GENETIC CODING. I. BIPOLAR BONDINGS IN THE TWO-STEP TAKEOVER OF PEPTIDE TEMPLATES BY PRENUCLEIC-RIBONUCLEIC ACIDS

BY

Gilbert TURIAN*

(Ms soumis le 18.6.1996, accepté le 4.11.1996)

ABSTRACT

Polarity at onset of genetic coding I. Bipolar bondings in the two-step takeover of peptide templates by prenucleic-ribonucleic acids. - A model of prebiotic evolution of coding molecules can be traced from the interplay of their weak versus strong bipolar bondings which link (1) the primordial 1–letter code of covalently bonded amino acids into (cyclo)peptides, through (2) its prenucleic takeover by 2–letter coding doublets of polyphosphor(P)amide(N)-bonded bases to (3) the "modern" 3–letter triplets of coding nucleotides singled-out by median insertion of ribose (P-R-N) at ribonucleic acid takeover.

Key-words: Peptide templates, prenucleic acid, ribonucleic acid, takeovers, polarity.

INTRODUCTION

The most primitive type of interatomic or intramolecular bonding is the covalent one which requests overcoming of bipolar electric repulsion to pair electrons (e^-) between two atoms, as first occurred in the bonding of two of the primordial atoms of hydrogen (H) into the first molecule, that of dihydrogen (H_2: H + H = H:H = H-H). Then, another landmark event of chemical evolution was the sharing of two electron pairs between the two H atoms and the O atom in the water, H_2O molecule.

It was of decisive consequence for the further prebiotic evolution that H_2O could combine its structural covalent bonding with a possibility of partial ionic bonding authorized by the breakage of one of its H to O covalent bonds. Such reversible ionization of H_2O which proceeds to only a very slight extent at standard temperature (25° C) and pressure produces an equilibrated proportion of a hydroxide anion (OH^-) and a hydrogen cation or proton (H^+, generally noted as the hydronium ion H_3O^+ by bonding to H_2O). Interestingly, in following the Bronsted-Lowry definition of a base as a proton acceptor and an acid as a proton donor, H_2O which accepts a proton as H_3O^+ is

* Laboratoire de Microbiologie générale, Université de Genève, 30, quai Ernest-Ansermet, CH-1211 Genève 4.

a base in the reaction while H^+ can be considered as the simplest, so-called Lewis acid. Thus, it is plausible that "all phenomena of the world are the reactions of one kind of acid with one kind of base" (ATKINS, 1991), if such reactions are considered in the realm of the unifying principle of electric bipolarity (TURIAN, 1989–1994).

The bipolar acid-base reactivity of H_2O was evolutionarily conserved in primordial organic molecules such as the biogenic amino acids built with covalently bonded backbone residues (C-C-N) but with ampholytic terminals conferring them pH-depending "zwitterion" properties at their isoelectric point. Consequently, only amino acids bearing heteropolar terminal groups (-COOH $^{(-)}$ *versus* -NH2 $^{(+)}$), and no other prebiotic molecules, such as dicarboxylic acids bearing only repulsive (-COOH) homopolar terminals, could insure polymeric chains formation. However, after such oligo-polypeptides are completed, only the two α-amino (+) and α-carboxyl (-) terminal groups continue to ionize and to present differences in their acid-base behavior and polarity at different pH values. All the other interchain groups of the constituent amino acids are covalently (*) joined in a repetitive pattern of peptide bonds (C-C-N*-C-C-N*-C-C- ...).

1. Peptide templates

a) Amide-bonded intrachain peptide sequences

A peptide chain is endowed with polarity; it starts at one end with an α-amino group that does not participate in peptide linkage while the terminal end of the chain contains an α-carboxyl group that is also free. Consequently, and on the prototypic model of covalently-structured H_2O molecules "inter-netted" by electrostatic interactions (ref. in TURIAN, 1995) the CO-NH peptide bonds between their constituent amino acids remain potentially electrically bipolar and endowed with reciprocal electrostatic affinities potentially generating interchain H bondings (1.b).

Chemically, the peptide bond is a substituted amide linkage between the anionic α-carboxyl (C) group and the cationic α-amino (N) group which removes one molecule of H_2O to close the covalent bond by the sharing of one pair of electrons between C and N atoms. However, at around neutral pH in pure H_2O, amino acids are not very reactive to polymerize because of their "zwitterion" form in which the negative charge on the carboxyl group (COO-) destroys the electrophilic character of its C atom while, simultaneously, the nucleophilic character of the N atom is destroyed by protonation (NH_3^+) of its pair of electrons. Therefore, to bond efficiently, amino acids need to be in an activated form provided either by the coupling of the COOH group with adenylate as well known for "modern" protein synthesis or by its condensation with linear-cyclic polyphosphates, dicyanimides or carbodiimides (see PONNAMPERUMA, 1978). In prevital simulated conditions, alanyl adenylate was made to polymerize in aqueous solutions in the presence of the clay montmorillonite (BRACK, 1976). Another primordial way to shift the thermodynamic equilibrium of the reaction toward the end-product (peptide) has been the separation of the other end product, H_2O, by heat (FOX & HARADA, 1960)

or the enrichment of the aqueous incubation solutions in anhydrizing agents (discussion in KAUFFMAN, 1993; role of Mg^{2+} in TURIAN, 1996).

The famous OPARIN's (1924, see 1957) postulate that life began with self-replicating proteins was widely supported by many other suggestions made by "origins of life" researchers and originating from that already made by LEDERBERG in 1961 that "there may be some primitive proteins which had the capacity of replicating". Among the earliest ones were those made by Fox since the 1950s and summarized in his 1965 book, in which he argued that synthesis of a small proportion of peptide bonds in aqueous solution is thermodynamically feasible without imparting energy to the system and that this aqueous system should be hypohydrous as provided by hydrolysable polyphosphates for linear polypeptides, Mg^{2+} ions for cyclic peptides (TURIAN, 1995) or, because of their ancient role in energy metabolism, thioesters in deep acidic waters (DE DUVE, 1991). Chemically, the energy barrier could also be surmounted by raising the temperature above the boiling point of water, or by using polyphosphoric acid which permitted polymerization of peptides at 70° C (HARADA & FOX, 1965).

The optimistic conclusions about an ordering of the thermal protoprotein polymers have, however, been tempered by JOYCE (1989) who commented that "the amino acid sequence of abiotic polypeptides would have been tremendously heterogenous" and, consequently, it remains difficult to estimate the possibility that a randomly produced polypeptide can catalyse peptide-bond formation with any significant degree of specificity. FERRIS (1989) also strongly argued against FOX's precepts that protocells were formed from such thermal polymers, even though he conceded that "it is not to say that polypeptides did not have a role in the first life, or that there might have some sequences of polypeptides that formed more readily than others". However, it does not seem likely that proteins with specific sequences formed without some sort of a catalyst to direct their ordered formation as, according to DE DUVE (1995) "there are good reasons to believe that relatively simple peptides would already be endowed with catalytic activities and thereby participated in protometabolism" (see also KAUFFMAN 1986).

All the above structuro-functional features of peptides somehow concur to our view that they were the earliest formed biogenic polymers, be that they originated by pairings between the primordial amino acids in OPARIN-HALDANE prebiotic broth ("primitive soup") successfully simulated in 1949 by Miller-Urey pioneering experiments, or were freed from carbonaceous meteorites. The first amino acids to enter the pregenetical code were the most abundant hydrophilic-hydrophobic glycine and alanine and possibly also those with anionic side groups (Asp, Glu or P-ser) which could be incorporated into the flat, polyanionic surface-bonded antiparallel β-sheets postulated by WÄCHTERSHÄUSER (1988).

Contrarily to the specific H bondings between nucleobases (G-C; A-U(T)), those between intrachain peptide bonds (-NH · · · OC-) have no proper informational content. Nevertheless, they play an indirect directional role in the self-recognition processes between opposite chain sequences. Such directive electrostatic interactions operate in an alternate way by inversely polarized H-bonds ordered by the opposite positionings of CO and NH groups, themselves inversed between the antiparallel chains of β-sheets.

Different kinds of constraints act on the conformation in space of an isolated polypeptide chain, among which the rigidity and *trans* configuration of the peptide bonds, the electrostatic repulsion or attraction between amino acid residues with charged R groups, the bulkiness of adjacent R groups. In such pairings, attraction of opposites versus self-attraction processes intervene. Most effective are attractions between amino acids with side groups which have opposite or complementary features. Such interactions between residues conform to PAULING's (1960) laws of chemical bonding which involve weak bondings by ionic bonds, hydrogen bonds in which an electropositive H atom is attracted to electronegative atoms such as O or N, van der Waals forces which arise from a nonspecific attractive force originating when two atoms come close to each other (H ··· H), hydrophobic forces between methyl groups, etc. H bonds also occur both in salt bridges with charged donor and acceptor groups and with uncharged groups (see in JEFFREY & SAENGER, 1994).

b) H-bonded interchain peptide sequences

Protein possesses not only primary structure, i.e., a covalent backbone, but also a characteristic secondary structure, the manner in which successive amino acid residues are arranged in space. Its α-helix and β-structures are stabilized by H bonds between peptide groups, intrachain in the case of the α-helix and interchain in the case of β-structure; these bonds, although individually relatively weak, collectively give such structures considerable stability. The β-pleated sheets can consist of only two strands in *parallel* or *antiparallel* orientation. However, since the sheet has free H-bonding groups at both edges, more strands can be added to form a multitude of strands.

In contrast to the α-helix which is formed by a continuous segment of a polypeptide chain, the β-pleated sheet is formed by different segments of the polypeptide. These sheets can coil to form cylinder-like structures referred to as β-barrels [which, as a ring, are "infinite" β-sheets] (see JEFFREY & SAENGER, 1994). However, these circularly folded segments can form either (a) pseudorings remaining connected by β-turns insured by vertical peptide bonds but leaving a possibility of free, weak H bondings at both edges of the turn or (b) complete rings individualized by horizontal, planar peptide bond closure of a circular polypeptide segment, thereby corresponding to a cyclic peptide; many such cyclic peptides, parallely or preferentially antiparallely oriented can be stacked by H bonds into nanotubes (GHADIRI *et al.*, 1993).

Cyclic peptides, when presynthesized by chemists or prebiotically assembled around mineral (Mg^{2+}-silicates, etc.) "nanostalagmites" as postulated (TURIAN, 1995) could serve, *in vitro* or in nature as templates, only for additional monomeric rings similar in their amino acids sequences. In contrast, pseudorings of β-sheets elongating by circular foldings of the polypeptide chain could easier be complementary in their amino acid units (opposed glu^- to lys^+, etc.) from one chain to the other. Two successive pseudorings could then close on themselves by H bondings at their edges thereby forming dimeric template rings (see Fig. 1 in TURIAN, 1995). However, such weakly closed

rings would keep a possibility of rupture at their double edge H bonding allowing reopening and further stretching as polypeptide segments. Such device would overcome the difficulty of copy of the cyclopeptide ring by normally strand-like nucleobase-polyphosphate chains by presenting relatively linearized amino acid sequences for their recognition at prenucleic coding takeover (see p. 221).

In the random polypeptide copolymers synthesized *in vitro*, there were some indications of the possible role of certain residues. Thus, the sequential polymer poly-L-alanylglycine synthesized by FRASER *et al.* (1965) formed β- sheets packed in pairs with the glycyl surfaces in contact. Such glycyl-glycyl as well as the alanyl-alanyl contacts suggest a positive interaction between their residues. In fact, it is known that the β-structures are favored by the presence of large numbers of glycine and alanine residues (WATSON, 1970). Such positive interactions could thus be expected in the interchain spaces of the stacked tetrapeptide rings possibly forming nanotubes after autocatalytic replication of the dipeptide (β-alanyl-glycine) units (TURIAN, 1996). Its oppositely positioned glycine as well as the CH_3- groups of two opposite alanines could be attracted by hydrophobic forces. The order of the amino acids (ala-gly-ala-gly) should thus be the same on the successive rings of the cyclopeptide. This would mean that homologous amino acids would face each other. Moreover, it is known that, in antibiotic polypeptide synthesis, amino acids are added in a predetermined sequence and that a built-in template is required that dictates addition in the right order (LIPMANN, 1971). There could also be predicted the existence of a specific surface on the peptide templates that attracts the amino acids and then lines them up in the correct order. Such templates or "attractors" would also have the capacity of serving as templates or "self-attractors" for themselves to self-replicate. In this process, their specific surfaces constituted by the side chains of successive amino acids would be, in someway, exactly copied to provide new templates.

In the rules of recognition and attraction that govern such selective binding of small molecules to their template, the attraction is based on relatively weak bonds and hydrophobic forces (see above, glycine-glycine, etc.) that can form without enzyme. Nevertheless, a chemical argument was made by WATSON (1970) against the existence of protein templates, according to whom "no template whose specificity depends upon the side groups of closely related amino acids, like glycine and alanine or valine and isoleucine, which both differ by only one methyl group, could ever be copied with the accuracy demanded for efficient existence, while only molecules with opposite or complementary features would preferentially attract each other". The answer to the question whether any copying process can be sufficiently accurate to distinguish between such closely related molecules might then depend upon the criterion of sufficient accuracy in the self-recognition of amino acids. In fact, this requirement could be considered as sufficient for their copy in the primitive conditions of prebiotic life, as suggested by the preferential recognition of homologous amino acids in the self-replication of the dipeptide bonds of a primitive cyclopeptidic template (TURIAN, 1996). Other peptides, but linear, had already been proposed by CALVIN (1969) to function as

templates for their self-replication. At last, since amino acids were among the most abundant biogenic building blocks available on the prebiotic Earth (MILLER, 1987), their spontaneous polymerization into linear or cyclic peptides would be "at least easier to visualize than the spontaneous assembly of RNA molecules" (DE DUVE, 1995).

2. Prenucleic acid takeovers

Considering the fact that evolutionary organization requires self-reproducing information storage, EIGEN (1995) has conceded that "RNA, or a precursor, would have been necessary to set the merry-go round of evolution in motion". On the more concrete ground of the known limitations in prebiotic syntheses (ribose, pyrimidines, etc.), PICCIRILLI (1995) has straightfully hypothesized that "the first kinds of nucleic acids might not have been quite the same as ours". Therefore, prenucleic acids would have had to takeover the approximate information encoded in the primordial peptide sequences (2a, b) before having been themselves taken over by true nucleic acids (§3).

The "protein first" evolutionary period provided the primeval templates of a pregenetical code of approximate accuracy but endowed with the selective advantage of amide-bonded esters of amino acid coding letter sequences not easily hydrolyzed and thus relatively resistant to the hard environmental conditions of prebiotic world. A selection pressure must have then developed in favor of template sequences better fitted to recognize mutually and pair code than amino acids.

Surface-bonded polyhemiacetal polymer structures with phosphoribose constituents (but ribose problem, see SHAPIRO, 1988) might have glycosidically bonded purines (WÄCHTERSHÄUSER, 1988). However, they do not appear to have provided the possibility of self-replicability of their base sequences requested by polymeric templates. Better fitted to assume that task would be polymers lining up coding nucleobases by (a) polyamide chains bondings (peptide nucleic acids PNA) or (b) phosphoramide bondings on polyphosphate chains.

a) Polyamide-bonded base chains

The recognition of an amino acid by a trinucleotidic codon is effected indirectly through a tRNA molecule, but recognition in pregenetical evolution may have occurred through direct amino acid–template interactions (WOESE *et al.*, 1966). The feasibility of interpolymer hydrogen bonding between the peptide backbone and nucleobases is thus pertinent to the evolutionary study of nucleic acid–protein interactions (NIU & BLACK, 1979, NIU *et al.*, 1987) in which a 5'-uracil substituent could recognize the side chain of a peptide-bound amino acid, and cytosine would have formed two hydrogen bonds to the peptide backbone plus one to the hydroxy-amino acids. The pregenetical material was not yet ribose-phosphate linked but peptide-linked. This noticeable principle was successfully followed in the synthesis described below of a new class of DNA and RNA analogs in which the sugar-phosphate backbone has been replaced by a similar linear structure of amino acids directly bonded to nucleobases.

Peptide nucleic acids or PNAs were first designed by computer modelling by NIELSEN *et al.* (1991) and synthesized by EGHOLM *et al.* in 1993 as a polymer in which the (deoxy)ribose-phosphate backbone is composed of N-(2-aminoethyl) glycine units to which nucleobases are attached as side chains by carbonyl methylene linkers. A polycytidine decamer of PNA acted as a template for oligomerization of activated guanosine mononucleotides in the presence of Na and Mg ions; nonetheless, the PNA template was not as efficient as the corresponding DNA template (BÖHLER *et al.*, 1995).

The known propensity of PNA molecules to pair with RNA and DNA according to the WATSON-CRICK base-pairing rules suggested a good starting point for the evolution of the possibility of pregenetic takeover. This molecular system can indeed serve as a template both for its own replication and for the formation of RNA from its subcomponents. Its molecular complexification to RNA would in principle be possible by the acquisition of the phosphodiester-α-N (3/1 or 9)-glycosyl bonds connecting D-ribose to the phosphate groups and the nucleobases (see HORGAN, 1996). However, it should neither be considered as "the primary replicator" nor that it even existed under prebiotic conditions (ORGEL, 1994).

b) Phosphoramide-bonded polyphosphate-base chains

In this takeover process, the pregenetic information present in a primordial peptide molecule would direct for the first time the order of specific nucleobases. This would necessarily implicate recognition of the mutual specific partners and a direct correspondance between certain amino acids and certain nucleobases, itself founded on preferential electrostatic affinities, hydrophobic–philic interactions, etc. Recognition by uracil side chains of specific hydrophilic amino acid side chains has been described since 1973 by BLACK. Even the number of atoms in an amino acid chain played a role in relation to pairs of coding bases and many other plausible arguments have been proposed on a physico-chemical basis for the establishment of the genetic code (JUNGCK, 1978, 1984).

The sparking event of this first phase of "coding evolution" could have been the progressively increased environmental proximity of newly synthesized nucleobases around amino acids lined up in preformed peptide sequences which would have led to their selective interactions directed by conformational affinities. Steric constraints around and bulkiness of the specific residues of amino acids would have restricted to couples (doublets) the number of bases habilitated to bilaterally "pinch" and weakly bound each selected amino acid as it is known that the 1st and 2nd letters of each "modern" codon are the primary determinants (JUNGCK, 1963 in 1978). This would necessarily mean that the formation of the first doublets of nucleobases was mediated by the physico-chemical affinities inherent to the amino acids-nucleobases partners. Therefore, such initially coding bases chosen by direct contact and recognition with their selected specific amino acid in fact correspond to those of "modern" anticodons rather than codons, those nucleobases which remained in the closest proximity to the translated amino acids. Such a view corresponds to that expressed by DUNNILL (1966)

who proposed relationships between the anticodon trinucleotides and amino acids thereby differing form PELC & WELTON (1966) who seeked to find a stereochemical relationship between amino acids and codons (leucine-UUG, etc. in their Fig. 1).

Consequently, and assuming that the base doublets first constituted by their specific trapping by amino acids have been perpetuated along the whole evolutionary span — with a more recent completion to triplets — then glycine would have been "caged" by two cytosines (C) and alanine by C + guanine (G) (Fig. 2). Such base doublets were in fact true original codons by retrograde reading and became anticodons only after their preservation by initial retrograde transcription (Fig. 1.a') into complementary base doublets -GG and GC-, the prenucleic precursors of the "modern" codons of mRNA in their now anterograde code reading.

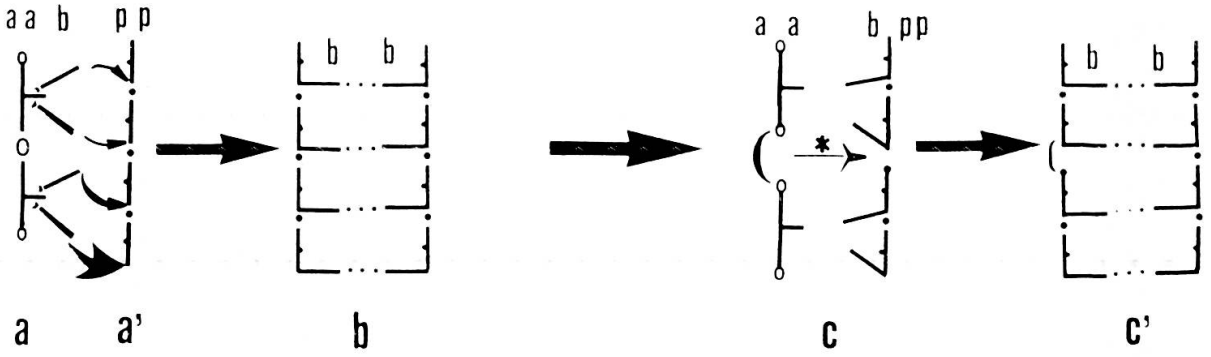

FIG. 1

Modellized polaro-informational back (a) and forth (c) coding transfers between peptide-bonded amino acids (aa) and phosphoramide-bonded nucleobases (bb) of prenucleic sequences. - Step a (retrotranslation): specific attraction by side chains of amino acids of nucleobase doublets (for their identity per amino acid, see Fig. 3) bilaterally "pinched" by weak bonds of their anterior pole(s); a') stabilization of the doublets by strong phosphoramide hooking of their N posterior pole on pyrophosphate (pp) units (details in Fig. 2) entraining a rupture by asymmetric tensional force of the apical weak bondings, and provoking a half-circle tilting over of the nucleobase-phosphate units promoted as prenucleic codons. - Step b (retrotranscription): complementation of the codons by their free apical H-bonding sites with specific partner bases also stowed on polyphosphate chains (double "strand" step). - Step c (anterotranslation): prenucleic codons now functioning as anticodons to "prey" on their specific amino acids, peptide-bonded ($-H_2O^*$) at the expense of the transiently hydrolysed inter-dipyrophosphate anhydride bonds before c') reconstitution of the polyphosphate-nucleobase continuity by dehydration condensation (c) and further "rematching" of the anticodons doublets (c') with nucleobases of the codon strand preserved since step b.

Naturally, after their initial retrotranslation from original amino acids (Fig. 1.a), the nucleobase doublets had to be immediately someway stabilized on polymeric chains. However, because of the temporary deficiency in adequate pentose sugars for reasons of poor and aspecific non enzymatic synthesis (see p. 224), the chains could not yet be diphosphoester-bonded D-ribose α, N-glycosylated base sequences (ribonucleic acids or RNAs, see 3) but, by necessity, polyphosphate chains only (Fig. 1.a'–c'). Following

KORNBERG's (1995) suggestion of a vicariant role of polyphosphates in the prebiotic evolution, we have relied on the possibility of a covalent amide bonding of nitrogenous bases on alternate phosphate groups, one per pyrophosphate unit of such prenucleic polymers (Figs 1 and 2).

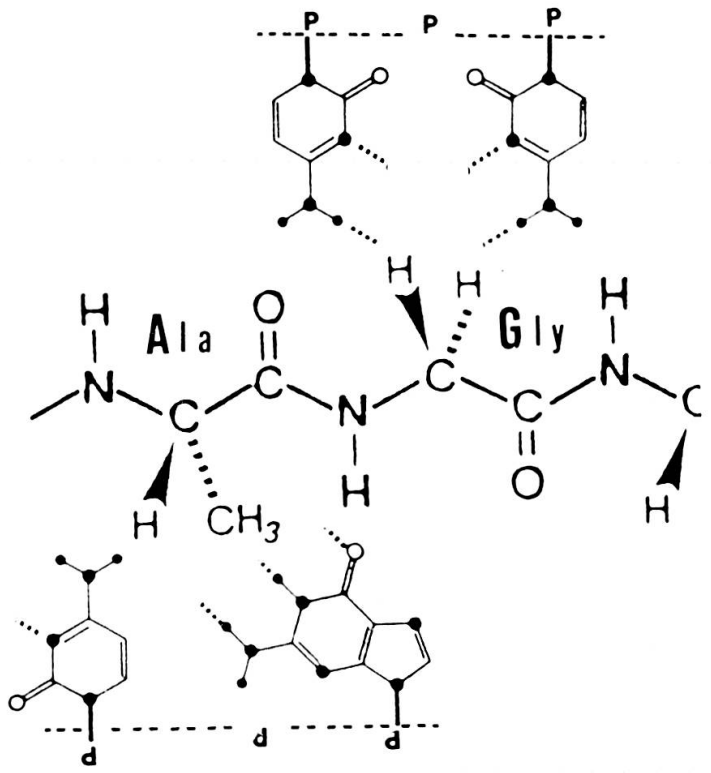

FIG. 2

Example of bilateral "pinching" of two selected amino acids — glycine + alanine — by the anterior poles of specifically attracted nucleobases preceeding the stabilization by N(1/9)-P bondings of the primordial coding doublets formed (= 'modern' anticodons).

A straight phosphoramide or N-P bond, the simplest occurring in $H_2N\text{-}PO_3H_2$, is more stable at alkaline than acid pH values as well-known in a biological phosphoramidate compound, creatine-phosphate or "phosphagen" energized (N~P) from ATP. Other phosphoramidates could be synthesized in neutral aqueous solutions on oligonucleotides by SHABAROVA (1988). SIEVERS and von KIEDROWSKI (1994) have used phosphoramide bondings as condensing agents for the replication of hexameric DNA oligomers while FERRIS *et al.* (1996) have used adenosine 5'-monophosphoramidate as precursor of the nucleoside 5'-phosphorimidazolide oligomerized on mineral surfaces. In the non-enzymatic conditions of our prebiotic scenario, we have postulated for the phosphoramidate syntheses a simple mechanism of driving a synthesis reaction beyond dehydration on the model proposed by FOX (1988) for peptide bond formation, namely a coupling of high-energy pyrophosphate (substitute for ATP) to nonspecific phosphorylation and dephos-

phorylation cycles. A tripolyphosphate dephosphorylation might thus be coupled with the dehydration intervening in the phosphoramido-bonding of the nucleobases ($triPO_3$ + nucleobase = $pyroPO_3$-base + H_3PO_4, in preparation).

The fundamental role of the polyphosphate backbone is not only to hook or stowe the nucleobases by their common N 1 or 9 atom. This atom occupies their posterior pole leaving the anterior pole free to weakly bind the affinity chosen aminoacids. It seems that such choices correspond to a certain logic, namely common C — the lightest base — for the simplest amino acids glycine (CC for -H), alanine (CG for $-CH_3$) to valine (CA for CH-$(CH_{3/2})$ but replaced by common A — the lightest purine — from valine on (Fig. 3). Such structuro-functional asymmetry would thus provide the mechano-chemical force for the necessary "elution" of the doublets from their anterior, transitory fixation pole. The resulting disequilibrated forces in favor of the basal stronger phosphoramide bond are then expected to drive an around 180° tilting-over movement

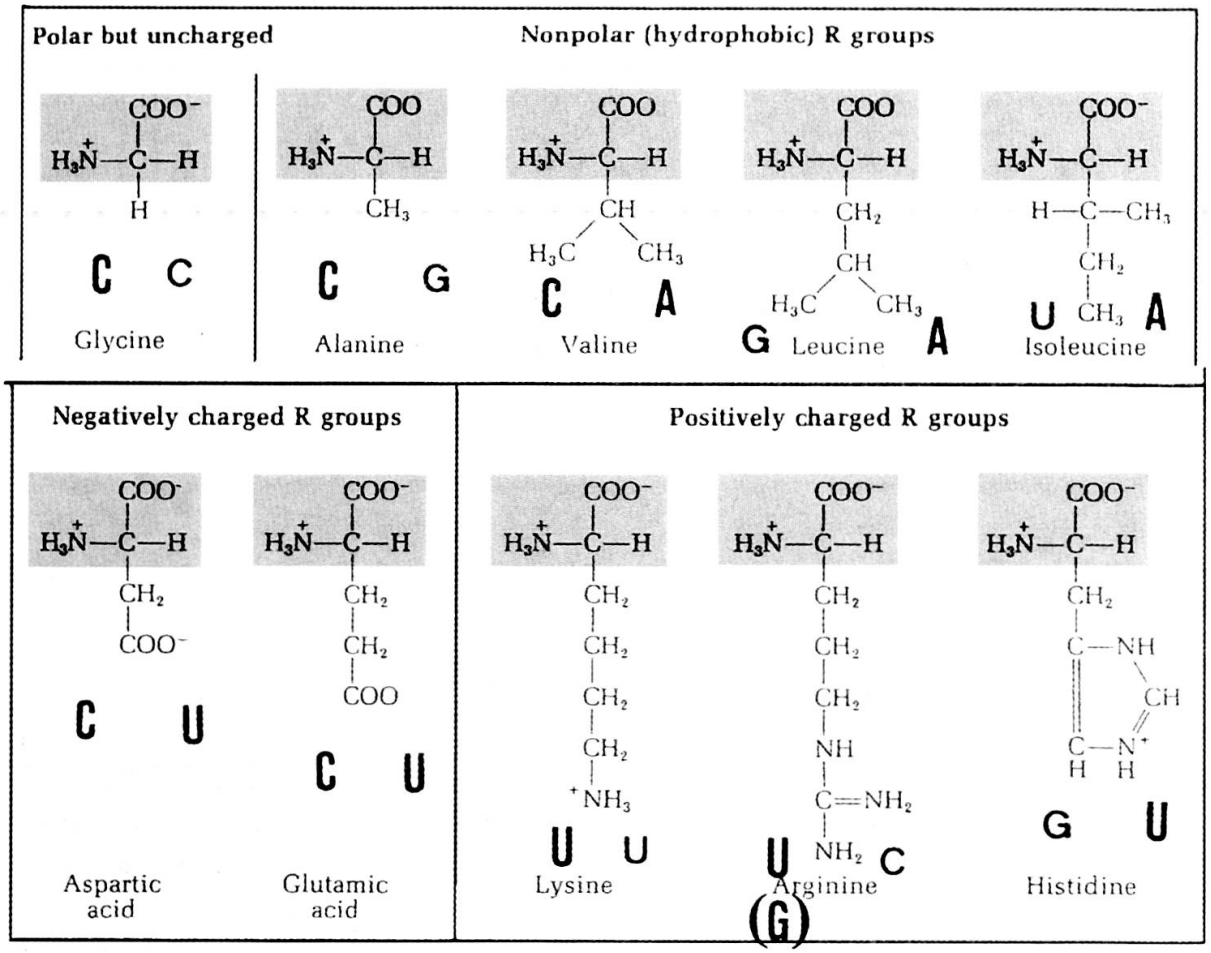

FIG. 3

Retrotranslation of the 1–letter primordial code by 2–letters (doublets) nucleobase codons corresponding to the first two bases of "modern" antero-translating anticodons. Group affinities of amino acids (polar-non polar, negative-positive R groups) for certain nucleobases expressed by one common (major capital) and one differential letter (minor capital).

of the doublets of nucleobases (Fig. 1.a'-b). Consequently, these bases, with their now free, potentially weak bonding anterior molecular poles, become available for H bonding with their conformationally complementary base partners (Fig. 1.b), those which remained such till "modern" genetic coding such as C-C inversed into G-G, C-G → G-C, U-A → A-U, etc. Those complementary bases would also be "captured", from their N1 pole for pyrimidines and N9 for purines, by phosphate groups which, by dehydration condensation, would in their turn also form complementary polynucleophosphate chains or strands which could insure preservation of the newly acquired coding sequences.

The newly formed codon strands would then be available to serve as templates for antero-transcription into anticodons competent to "predate" on free amino acids (Fig. 1.c) to line them up in the correct order, thereby reproducing reliably the primitive peptides previously self-replicating more approximately. This choice of the amino acids should be determined by their correlative properties with those of their respective anticodons. Thus, it is presumed that the doublet C-C will "cage" glycine while C-G will shelter alanine. The loss of 1 molecule of H_2O then entrained by the peptide bond formation between the two successive amino acids could be driven by the parallel opening, by hydration, of the phosphate bond linking two successive base doublets in the polyphosphate chain. Parallely, this transient dislocation mechanism would rupture the weak apical bonding between the bases and their elected amino acid so that this "elution" will liberate the doublets-phosphates which would only be reconstituted as a continuous anticodon chain by dehydration condensation oppositely to a complementary template codon chain (Fig. 1.b).

A next round of translation will be insured by renewed rupture of the interchain weak H bonds (Fig. 1.c), possibly by some environmental effector such as heat. Presumably, this separation of the polynucleophosphate strands would be facilitated by the fact that, as the pyranosyl RNA strands (see p. 224), they may not twist around each other, a useful feature because in a world still without protein enzymes (topoisomerases), twisting could prevent the strands from separating cleanly on preparation for replication.

3. Ribonucleic acid takeover

This second takeover involves two major innovations, namely molecular with the transition "nucleobases – nucleosides – nucleotides" which difficultly incorporated ribose between a base and two phosphate groups, and operational with the completion of the "archetypal" doublets of coding bases by a third base. Such transitions and their evolutionary timing are still poorly understood. It is therefore of no surprise to remark that practically all models of genetic coding proposed to date involve ribonucleotides – thereby eluding the ribose problem. As such they are already fully part of the "RNA world" (GILBERT, 1986) born with the discovery of ribozymes by CECH and ALTMAN in the 1970's (see GESTELAND & ATKINS, 1993), namely a RNA system which could be endowed with both informational and catalytic properties. Some of the newly synthe-

sized ribozymes (WILSON & SZOSTAK, 1995) which do not involve only the ribosome-phosphate backbone but also the bonding between carbon and nitrogen in polyamide bonds such as in PNA (2.b) might, according to SZOSTAK (see HORGAN, 1996), be "part of a larger class of versatile molecules produced in nature".

The central question of how self-replicating RNA was evolutionarily produced thus remains unsolved when we consider the difficulties of the synthesis of one of its key components, D-ribose. Its prebiotic precursor, formaldehyde, can oligomerize by formose reaction to a mixture of sugars from which only low concentrations of D-ribose could be resolved (see CAIRNS-SMITH, 1982; JOYCE, 1989, 1992). To add to the above difficulties, bases and ribose molecules must then come together to react and make nucleosides. On the nitrogenous side, if prebiotic simulated experiments have easily yielded purine nucleosides, this was not the case with pyrimidine nucleosides which could be formed but only with lower efficiency (see however ROBERTSON & MILLER, 1995), thereby reenforcing the opinion that it is unlikely that an abundant pool of RNA building-blocks existed on the primitive Earth. In his surface-bonded ribonucleotides theory involving "phosphoribose" polymers, WÄCHTERSHÄUSER (1988) considered that "the presence of ribose in nucleic acids poses a major problem for all versions of the prebiotic broth theory" and also mentioned ribose precursors such as phosphotrioses to be rather uneffective substitutes of ribose. In this line, ESCHENMOSER, as quoted by ORGEL (1994) circumviated the problem by prospecting oligonucleotides containing a 6-member sugar ring, pyranose. The pyranosyl RNA strands built were efficient to combine into double-strands units but they did not twist around each other as they do in double-strand RNA.

As for replication of RNA sequences in the absence of enzyme, it has met with limited success (JOYCE, 1989, 1991; JOYCE & ORGEL, 1993; ORGEL 1994) and experimental synthesis of ribonucleotides was difficult to achieve all-the-more that they preferentially link by 2'-5' bonds rather than the biologically requisited 3'-5' bonds (ORGEL, 1992). Consequently, life may rather have begun with some simpler replicating system such as that of our proposed sugarless prenucleic acid.

DISCUSSION

Life is no longer considered as a magical force but as an emergent property based on the behavior of complexified organic polymers self-assembled from elementary molecular building blocks along three major evolutive steps of genetic coding: (1) single peptidic amino acids, (2) double prenucleic nucleobase-phosphates, (3) triple nucleobase-sugar-phosphate monomeric units of nucleic acids. Primeval peptide sequences could have mainly resulted from non enzymatic anhydrizing condensations of amino acids in relatively alkaline aqueous media (TURIAN, 1995, 1996), acid-requiring thioester-mediated bondings (DE DUVE, 1991) remaining confined to specialized ecological sites (deep marine holes, etc.). The 2nd step of prenucleic takeover by phosphoramide-bonded

nucleobase (N)-phosphates (P) endowed with acid sensitivity could therefore have imprinted to the whole scenario of life's origins its requirement for fairly neutral (pH 6.5–7.5) conditions. These could have been provided by watery media layered on the ruguous surface of primeval rocks, a surface requirement first suggested in 1950 by BERNAL (in FOX, 1965), then by WÄCHTERSHÄUSER (1988), and recently by FERRIS *et al.* (1996). The ribose insertion singling-out the 2rd step of ribonucleic takeover would have more recently provided a selective advantage over the prenucleic coding units by the increased stability of RNAs over a wider range of pH, thereby securing the necessary coding accuracy.

To insure the onset of genetic coding, two major biochemical problems were to be overcome, namely acquisition of a ribose-phosphate backbone and specific ordering of nitrogenous bases on the newly synthesized D-ribose molecules. As for the PNA type of prenucleic acid (2.a), a tentative answer to the first question has been that "polyamide system might have folded into catalysts that generated D-ribonucleotides" (PICCIRILLI, 1995). In contrast, the molecular topology of the prenucleic type of polynucleophosphate chains (2.b) presents two more convenient possibilities of RNA takeover because of (1) facilitated intercalation of ribose on the middle of the monomer (Fig. 1.a') by phosphoramide bond acid hydrolysis permitting N(1/9)-glycosyl ester bonding with the nucleobase, on one side, and phosphoester bonding, on the other; (2) maintenance by the polyphosphate chain of the preceedingly acquired coding sequentiation of base doublets specific for given amino acids.

In conclusion, whatever pertinent above proposals may be, our understanding of the transition from pregenetic systems to an RNA world still deserves JOYCE's (1989) comments that "it is fraught with uncertainty and plagued by lack of relevant experimental data".

We are grateful to Ariane Fehr for typing the manuscript.

RÉSUMÉ

Polarité à l'origine du codage génétique. I. Liaisons bipolaires impliquées dans la relève en deux étapes des moules peptidiques par les acides prénucléique et ribonucléique. - Un modèle d'évolution prébiotique de molécules codantes reposant sur l'interaction de leurs liaisons faibles et fortes peut relier (1) le code originel à 1 lettre d'acides aminés liés par covalence en (cyclo)peptides, par (2) sa relève par code prénucléique à doublets de 2 lettres de bases associées par liaisons polyphospho (P)-amidiques (N), au (3) code moderne à 3 lettres par triplets de nucléotides, singularisé par l'insertion médiane du ribose (P-r-N) à la relève par l'ARN.

REFERENCES

ATKINS, P.W. 1991. *Atoms, Electrons and Change*, W.H. Freeman, New York, 263 pp.

BLACK, S. 1973. *Adv. Enzymol. Related Areas Mol. Biol.* 38: 193–234.

BÖHLER, C., P.E. NIELSEN & L.E. ORGEL. 1995. Template switching between PNA and RNA oligonucleotides. *Nature* 376: 578–581.

BRACK, A. 1976. Polymérisation en phase aqueuse d'acides aminés sur des argiles. *Clay Minerals* 11: 117–120.

CAIRNS-SMITH, A.G. 1982. *The Genetic Takeover and the Mineral Origins of Life*. Cambridge Univ. Press. Cambridge. 477 pp.

CALVIN, M. 1969. *Chemical Evolution*. Oxford University Press (Clarendon).

CRICK, F.H.C. 1968. The Origin of the Genetic Code. *J. Mol. Biol.* 38: 367–379.

DE DUVE, C. 1991. *Blueprint for a Cell: The Nature and Origin of Life*. Neil Patterson Publ., Carolina Biol. Supply Co., Burlington, North Carolina. 300 pp.

DE DUVE, C. 1993. The RNA world: before and after? *Gene* 135: 29–31.

DE DUVE, C. 1995. *RNA without Protein or Protein without RNA?* Chapt. 7 *in*: *What is Life? The Next Fifty Years*. Edit. MURPHY, M.P. & L.A.J. O'NEILL, Cambridge University Press. 191 pp.

DUNNILL, P. 1966. Triplet nucleotide-amino-acid pairing; a stereochemical basis for the division between protein and non-protein amino-acids. *Nature* 210: 1267–1268.

EGHOLM, M. & 9 Coll. 1993. PNA hybridizes to complementary Oligonucleotides obeying the Watson-Crick Hydrogen-bonding rules. *Nature* 365: 566–568.

EIGEN, M. 1995. *What will Endure of 20th Century Biology?* Chapt. 2 *in*: *What is Life? The Next Fifty Years*. Edit. MURPHY, M.P. & L.A.J. O'NEILL, Cambridge Univ. Press, CAMBRIDGE: L)L PP:

ESCHENMOSER, A. & E. LOEWENTHAL. 1992. *Chem. Soc. Rev.* 21: 1–16 (see ORGEL, 1994).

FERRIS, J.P. 1989. Issues of organic origins. Book Review. *Nature* 337: 609–610.

FERRIS, J.P., A.R. HILL JR, R. LIU & L.E. ORGEL. 1996. Synthesis of long prebiotic oligomers on mineral surfaces. *Nature* 381: 59–61.

FOX, R.X. 1988. *Energy and the Evolution of Life*. W.H. Freeman, New York.

FOX, S.W. 1965. *In*: *The Origins of Prebiological Systems and of their Molecular Matrices*. Pp. 361–382. S.W. Fox (Ed.), Academic Press, New York, 400 pp.

FOX, S.W. & K. HARADA. 1960. *J. Amer. Chem. Soc.* 82: 37–45 (see in FOX, 1965).

FRASER, R.D.B., T.P. MACRAE, F.H.C. STEWART & E. SUZUKI. 1965. Poly-L-alanylglycine. *J. Mol. Biol.* 11: 706–712.

GESTELAND, R.F. & J.F. ATKINS (Eds). 1993. *The RNA World*. Cold Spring Harbor Laboratory Press.

GHADIRI, M.R., J.R. GRANJA, R.A. MILLIGAN, D.E. MCREE & N. KHAZANOVICH. 1993. Self-assembling organic nanotubes based on a cyclic peptide architecture. *Nature* 366: 324–327.

GILBERT, W. 1986. The RNA world. *Nature* 319: 618.

HARADA, K. & S.W. FOX. 1965. The thermal synthesis of amino acids from a hypothetically primitive terrestrial atmosphere. Pp. 187–201 in S.W. FOX (Ed.).

HORGAN, J. 1996. The world according to RNA. *Scient. Amer.* 274(1): 16–17.

JEFFREY, G.A. & W. SAENGER. 1994. *Hydrogen Bonding in Biological Structures*. Springer-Verlag, New York. 569 pp.

JOYCE, G.F. 1989. RNA evolution and the origins of life. *Nature* 338: 217–228.

JOYCE, G.F. 1991. The rise and fall of the RNA world. *New Biologist* 3: 399–407.

JOYCE, G.F. 1992. Directed molecular evolution. *Scient. Amer.* 267(6): 48–55.

JOYCE, G.F. & L.E. ORGEL. 1993. In *The RNA World*. Pp. 1–25. GESTELAND & ATKINS Eds.

JUNGCK, J.R. 1978. The genetic code as a periodic table. *J. Mol. Evol.* 11: 211–224.

JUNGCK, J.R. 1984. The adaptationist program in molecular evolution: the origins of genetic codes. *Mol. Evol. Protobiol.* 345–364 (Ed. Matsunok, K. Dose & K. Harada). Plenum Press, New York.

JUNGCK, J.R. & S.W. FOX. 1973. Synthesis of oligonucleotides by proteinoid microspheres acting on ATP. *Naturwiss.* 60: 425.

KAUFFMAN, S.A. 1986. Autocatalytic sets of proteins. *J. Theor. Biol.* 119: 1–14.

KAUFFMAN, S.A. 1993. *The Origins of Order. Self-Organization and Selection in Evolution.* Oxford Univ. Press. 709 pp.

KORNBERG, A. 1995. Inorganic polyphosphate: toward making a forgotten polymer unforgettable. *J. Bacteriol.* 177: 491–496.

LIPMANN, F. 1971. Attempts to map a process evolution of peptide biosynthesis. *Science* 173: 875–884.

MILLER, S.L. 1987. Which organic compounds could have occurred on the prebiotic Earth? *Cold Spring Harbor Symposia on Quantitative Biology* LII: 17–27.

NIELSEN, P.E., M. EGHOLM, R.H. BERG & O. BUCHARDT. 1991. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* 254: 1497–1500.

NIU, C.-H. & S. BLACK. 1979. Hydrogen bonding in solution between the uracil ring and the peptide backbone demonstrated by nuclear magnetic resonance spectroscopy. *J. Biol. Chem.* 254: 265–267.

NIU, C.-H., K.-H. HAN, H.J.C. YEH & S. BLACK. 1987. Hydrogen bonding between cytosine and peptides of threonine or serine: is it relevant to the origin of the genetic code? *Biochem. Biophys. Res. Comm.* 14: 456–462.

OPARIN, A.I. 1957. *The Origin of Life on Earth.* Academic Press, New York.

ORGEL, L.E. 1992. Molecular replication. *Nature* 358: 203–209.

ORGEL, L.E. 1994. The origin of life on the Earth. *Scient. Amer.* 271(4): 53–61.

PAULING, L. 1960. *In*: LEHNINGER, A.L. 1982. *Principles of Biochemistry.* Worth Publ., Inc. New York.

PELC, S.R. & WELTON, M.G.E. 1966. Stereochemical relationship between coding triplets and amino-acids. *Nature* 20: 868–872.

PICCIRILLI, J.A. 1995. Origin of Life: RNA seeks its maker. *Nature* 376: 548.

PONNAMPERUMA, C. 1978. Prebiotic molecular evolution. In: *Origin of Life.* ed. H. Noda, pp. 67–81. Tokyo: Center for Academic Publ. Japan.

ROBERTSON, M.P. & S.L. MILLER. 1995. An efficient prebiotic synthesis of cytosine and uracil. *Nature* 375: 772–773.

SHABAROVA, Z.A. 1988. Chemical development in the design of oligonucleotide probes for binding to DNA and RNA. *Biochimie* 70: 1323–1334.

SHAPIRO, R. 1988. Prebiotic ribose synthesis: a critical analysis. *Origins of Life and Evol. Biosphere* 18: 71–85.

SIEVERS, D. & G. VON KIEDROWSKI. 1994. Self-replication of complementary nucleotide-based oligomers. *Nature* 369: 221–224.

TURIAN, G. 1989–1992. Polarity. From dipoles to biopolarizations. *Archs Sci. Genève* 42: 1–323. Addenda I–II, pp. 324–529.

TURIAN, G. 1994. *Polarity. From Electromagnetic Origins to Biological Take-over.* Verlag Dr. Kovac, Hamburg, 522 pp.

TURIAN, G. 1995. New Trends in Polarity. III. Dipolar hydrogen bondings as homotemplate forces for pregenetical evolution. *Archs Sci. Genève* 48: 173–182.

TURIAN, G. 1996. Autocatalytic replication of the dipeptide bondings of a cyclo-tetrapeptide enforced by Mg2+-saline anhydrization. *Archs Sci. Genève* 49: 45–50.

WÄCHERSHÄUSER, G. 1988. Before enzymes and templates: theory of surface metabolism. *Microb. Reviews* 52: 452–484.

WATSON, J.D. 1970. *Molecular Biology of the Gene.* 2nd ed. W.A. Benjamin, Inc. New York, 662 pp.

WILSON, C. & J.W. SZOSTAK. 1995. *In vitro* evolution of a self-alkylating ribozyme. *Nature* 374: 777–782.

WOESE, C.R. *et al.* 1966. In 1972. *The Genetic Code: The Molecular Basis for Genetic Expression.* Harper & Row, New York.