ORIGIN OF THE GENETIC CODE: A PHYSICAL-CHEMICAL MODEL OF PRIMITIVE CODON ASSIGNMENTS

JOSEPH NAGYVARY

Dept. of Biochemistry and Biophysics, Texas A & M University, College Station, Tex. 77843, U.S.A.

and

JANOS H. FENDLER

Dept. of Chemistry, Texas A&M University, College Station, Tex. 77843, U.S.A.

Abstract. Selective compartmentalization of amino acids and nucleotides according to their polarities is proposed as a physical-chemical model for the origin of the genetic code. Assumptions made in this hypothesis are: (1) an oil-slick covered the surface of the primitive ocean, constituents of which formed association colloids or micelles at the water-oil-air interfaces; (2) depending on the polarity of the media, these aggregates possessed hydrophilic and hydrophobic interiors where selective uptake of amino acids and nucleic acid constituents could take place; and (3) condensation and polymerization in the micellar phase were enhanced. According to the chromatographically observed polarities, for example, lysine and uridylate fall into the hydrophilic compartment, and phenylalanine and adenylate are enriched in the hydrophobic environment. These components could eventually be condensed to form a charged adaptor loop with an anticodon which is complementary to the presently valid codon. Only two groups of amino acids, hydrophilic and hydrophobic, were recognized by the primitive translation mechanism. Implications of this hypothesis for the further development of the genetic code is discussed. The catalytic power of micelles have been substantiated by successful synthesis of nucleotides under relatively mild conditions using thiophosphates as high energy phosphates.

No rational interpretation has been advanced to-date for the observed codon assignments (Woese *et al.*, 1966; Orgel, 1968; Crick, 1968). Direct interactions of amino acids to oligonucleotides which would have been most satisfying could not be found (Woese *et al.*, 1966). The reported codonic alignment of AMP on polylysine is probably due to a simple effect of the polycation (Lacey and Pruitt, 1969). One of the meaningful experimental approaches, the polynucleotide-directed incorporation of amino acids into protenoid microspheres is not free of ambiguity and requires itself an interpretation (Nakashima and Fox, 1972). It is generally assumed that the code in its most primitive form could only differentiate between two classes of amino acids, i.e. hydrophilic and hydrophobic (Orgel, 1968; Crick, 1968; Orgel, 1972). The evolution of the precise assignments could have been accidental (Crick, 1968). The grouping of codons and amino acids by similar polarity criteria has been advocated (Woese *et al.*, 1966).

Here we propose a physical-chemical model involving the selective uptake of amino acids and nucleotides in compartments composed of simple, and progressively more complex, association colloids. Subsequent to compartmentalization, polymerization and condensation could be enhanced in the microenvironment of such colloidal aggregates. This hypothesis is based on the suggested composition of the primordial ocean

Origins of Life **5** (1974) 357–362.

(Lasaga *et al.*, 1971) and on the recognized properties of relatively simple synthetic and biologically occurring aqueous, or 'normal', and nonaqueous, or reversed micelles and selective substrate solubilization therein (Fendler and Fendler, 1970; Fendler *et al.*, 1973). The proposed oilslick of 1 to 10 m thickness on the surface of the primitive ocean presumably predominantly consisted of saturated and unsaturated hydrocarbons and contained amines, bicarbonates, nitriles, carboxylic acids, and heterocyclic compounds in low concentration which were capable of forming, and/or reacting with hydrocarbons to form, association colloids and hence micelles. Advantages of our proposal are that

(1) it includes some of the essential features of the present code,
(2) rules out accidental allocation of amino acids to codons, and
(3) most significantly, it could be approached experimentally.

In dilute solutions in the absence of enzymes or other suitable templates, the simplest way to ensure the charging of the primitive adaptor with the amino acid is the compartmentalization of these components. Neglecting the actual mechanism of amino acid and nucleotide condensation for the time being, if a compartment contains only nonpolar amino acids such as phenylalanine and leucine and mostly adenylate, this system can lead to the predominant synthesis of a charged adaptor with an A-A-A anticodon (loop). The same holds for some other amino acids and their respective anticodons. Assuming simple partitioning, the amino acids and nucleotides would be distributed on the basis of their solubilities, the order of which is akin to that of the chromatographic values in a similar system. Accordingly, phenylalanine and adenylate would be enriched in the oil phase while uridylate and lysine would remain in the water phase. Rudiments of nucleotide and amino acid alignments according to their polarity behavior have been suggested by Woese previously (Woese *et al.*, 1966). Such idea has not found fruition until now, primarily because the amino acids and codons were assumed to be of the same polarity instead of the more propitious grouping of amino acids with anticodons. Some of the assigned polarities of nucleotides are in fact in error (Woese *et al.*, 1966). On theoretical grounds uridylate was listed, for example, as less polar than adenylate, whereas chromatography establishes the reverse order (Wyatt, 1955; Lohrmann *et al.*, 1966) (Table I). Simple distribution of amino acids and nucleotides in the aqueous and non-aqueous phases according to their polarities is clearly insufficient to bring about the desired interactions since the concentration of nucleic acid constituents would be undoubtedly very low in each phase. Compartmentalization of substrates in relatively high concentration in the micellar pseudophases renders such interactions highly feasible. Formation, structure and physical chemical properties of simple aqueous and non-aqueous micelles as well as substrate solubilization are well established (Fendler and Fendler, 1970; Fendler *et al.*, 1973). Reversed micelles contain a hydrophilic cavity which can solubilize polar compounds such as uridylate and lysine. Less polar amino acids and nucleotides, on the other hand, are likely to be enriched in aqueous micelles. The selectivity for the nucleotide monomers would probably be surpassed by an even stronger discrimination at the level of dimers and trimers and their amino acid anhydrides. Since a

TABLE I

R_F values of nucleotides

	Solvent	
	Isobutyrate-H_2O	Phenol-H_2O
Adenosine-2′-phosphate	0.49	0.70
Guanosine-2′-phosphate	0.24	0.46
Uridylic acid	0.24	0.35
Cytidylic acid	0.37	0.57
U-U-U	0.42	
A-A-A-	2.22	
C-C-C	1.45	
G-G-G	0.25	

All data taken from Wyat (1955) and Lohrmann *et al.* (1966).

variety of different types of micelles is feasible with differing surfactant structure and medium polarity, the uneven distribution of all existing amino acids and nucleotides seems possible. The principle of *similis simili gaudet* would govern the aggregation of micelles and their captive ingredients as well. Although our discussion will be restricted to the A-U system, the enrichment of G, I and C can also be envisioned. G and I could be taken up into particular micelles, especially in the presence of aromatic compounds.

At this point one should remember the significant differences between the individual amino acids and nucleotides with respect to their chromatographic mobilities (Woese *et al.*, 1966). The separation of tRNAs via chromatography and counter-current distribution is even more remarkable if their similar size is considered. The numerous substituents such as acyl, alkyl and thio groups occurring in tRNAs without much apparent functional justification could have had their origins as discriminators with respect to partitioning.

The postulation of micelles as the crucial loci in prebiotic evolution has other merits apart from the opportunity of compartmentalization. Micelles have catalytic powers which at times approximate those of enzymes (Fendler *et al.*, 1973). The relative exclusion of water could have facilitated the condensation reactions which belong to the least understood chapter of prebiotic chemistry. The size of the intra-micellar cavity could have been a decisive factor in determining the size of the first adaptors. Additionally, the protection of labile compounds against hydrolysis by reversed micelles could have functioned as a further selective force.

For our present discussion it is not very critical to elaborate on how and where the condensation of nucleotides with each other and with amino acids could have taken place. For instance, it is also possible that small oligomers were first formed in the aqueous phase and subsequently compartmentalized for further modification. Elsewhere, we shall argue that thiophosphates and thionucleotides could have played an important role in the process. The intermediary formation of aminoacyl nucleotide anhydrides, one of which was first proposed by Paecht-Horowitz *et al.* (1970), is

particularly attractive because such compounds could be polymerized to oligonucleotides bearing an activated amino acid.

The evolution of the ideal system could have progressed rapidly from relatively simple micelles through larger ones to the coacervates of Oparin (1965) or to the microspheres of Fox (Nakashima and Foe, 1972). The prebiotic random synthesis of macromolecules must have preceded this step. The crudest form of primitive cell contained, in our opinion, both aqueous and reversed micelles in its 'cytoplasm', where the polynucleotide template eventually associated with the various charged adaptors. Figure 1 illustrates such a primitive system of translation machinery. The adaptors could have been simple loops held together by A-U base pairs. In addition to the Phe-AAA adaptor, which is shown, the nonpolar compartment could have also contained the anticodons AAU, AUA, and UAA in different adaptors, also charged by leucine, isoleucine and tyrosine. Conversely, the adaptors in the reverse micelle could have possessed UUU and AUU in their loops and lysine and asparagine at their end. One may extend this model to other nucleotides and amino acids. Glycine and cyti-

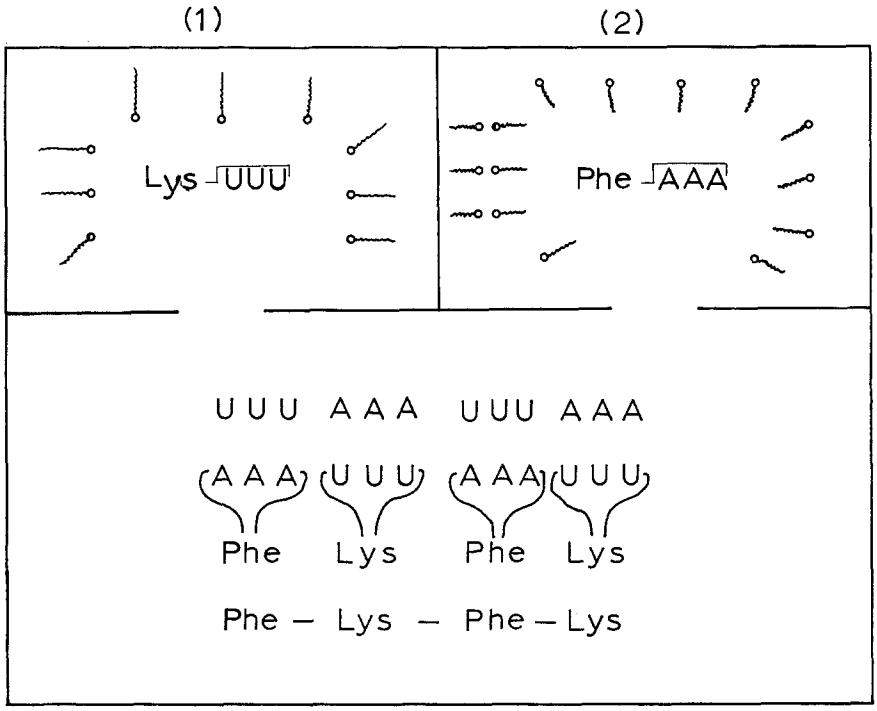

Fig. 1. Scheme of primitive translation machinery. Compartment 1 contains polar cavities which can be provided by reversed micelles and/or membranelike aggregates. Compartment 2 possesses a hydrophobic environment. Compartment 1 is capable of enriching lysine, uridylate and their short oligomers which are eventually condensed with some adenylate to form the adapter complex. The symbol Lys–⌐UUU⌉ stands for as mall loop which is held together by A-U pairs. Phe–⌐AAA⌉ may be formed in compartment 2 in a similar fashion. The arbitrary template **5** UUU-AAA-UUU - would be translated to Phe-Lys-Phe.

dylate, and proline and inosinate, respectively, could have fallen into the same compartments. The polypeptide which was eventually produced on the template should have reflected the composition of the message, albeit not necessarily with great fidelity. For the evolution of more discrimination adaptors the peptide products must have first enlarged and differentiated the micellar compartments, and then, eventually, developed a specific compartment of their own, i.e. recognition site for aminoacyl-adenylate and tRNA. The capacity of the adaptors to fit in a geometrical sense into compartments has continuously evolved, and so did the ability of the anticodon loop to discriminate, starting from a simple partitioning and culminating in an accurate recognition of a molecular environment. It is amazing that even in the present tRNAs the sum of hydrophobic and hydrophilic interactions, as evidenced in countercurrent distribution mobilities, reflects the composition of the anticodon (Woese *et al.*, 1966).

This hypothesis is intended to be, first of all, an interpretation of some basic codon assignments. It suggests that the amino acid-codon relationship, at least with respect to polarity, is not accidental, but it is the result of seemingly trivial physical-chemical factors which originally did not possess the criteria of direct recognition. We suggest that the accurate and specific ligases evolved from simple association colloids in a less obvious manner in accordance with the Principle of Continuity. The asymmetric peptides proposed by Orgel (1972), or some of the proteinoids prepared by Fox (Nakashima and Fox, 1972) may contain regions of discriminatory tendencies in the sense as suggested above. Because of their simplicity and similarity to biological membranes reversed micelles would offer a first approximation and obvious starting point for the study of primitive adaptors.

Experimental verification of all aspects of this hypothesis may well be time-consuming, but it is quite feasible. It is necessary to demonstrate that all condensation reactions, *viz.* phosphorylation of nucleosides, formation of aminoacyl nucleotides, and oligonucleotides, can be accelerated by colloid systems. We are currently using the energy of thiophosphate and S-cyanoethyl phosphorothioate to effect such condensations. So far we have found that all ribo- and deoxyribonucleosides can be converted to natural nucleotides at 60° and 70° in organic solvents; the 5′ nucleotides are the dominant products. Preliminary experiments also show that nucleoside phosphorothioates can be polymerized to form natural oligonucleotides. Selective uptake of polar nucleotides by reversed micelles has also been demonstrated in our laboratory.

Acknowledgement

This work was supported by the Robert A. Welch Foundation.

References

Crick, F. H. C.: 1968, *J. Mol. Biol.* **38**, 367.
Fendler, E. J. and Fendler, J. H.: 1970, *Adv. Phys. Org. Chem.* **8**, 271.
Fendler, E. J., Chang, S. A., Fendler, J. H., Medary, R. T., El Seoud, O.A., and Woods, V. A.: 1973, in E. H. Cordes (ed.), *Reaction Kinetics in Micelles*, Plenum Press, New York, p. 127.

Lacey, J. C. and Pruitt, K. M.: 1969, *Nature* **233**, 799.
Lasaga, A. C., Holland, H. D., and Dwyer, M. J.: 1971, *Science* **174**, 53.
Lohrmann, R., Soll, D., Hayatsu, H., Ohtsuka, E., and Khorana, H. G.: 1966, *J. Amer. Chem. Soc.* **88**, 819.
Nakashima, T. and Fox, S. W.: 1972, *Proc. Nat. Acad. Sic. (U.S.A.)* **69**, 106.
Oparin, A. I.: 1965, in S. W. Fox, (ed.), *The Origins of Prebiological Systems*, Vol. I, Academic Press Inc., New York, p. 331.
Orgel, L. E.: 1968, *J. Mol. Biol.* **38**, 381.
Orgel, L. E.: 1972, *Israel J. Chem.* **10**, 287.
Paecht-Horowitz, M., Berger, J., and Katchalsky, A.: 1970, *Nature* **228**, 636.
Woese, C. R., Dugre, D. H., Dugre, S. A., Kondo, M., and Saxinger, W. C.: 1966, *Cold Spring Harbor Symp. Quant. Biol.* **31**, 723.
Wyatt, G. R.: 1955, in E. Chargaff and J. N. Davidson (eds.), *The Nucleic Acids*, Academic Press Inc., New York, p. 256.