THE POSSIBLE ROLE OF ASSIGNMENT CATALYSTS
IN THE ORIGIN OF THE GENETIC CODE

VAHE BEDIAN*
Department of Biophysical Sciences, State University of New York at Buffalo, Buffalo, New York, 14214, U.S.A.

(Received 29 January, 1982; Accepted 5 April, 1982)

Abstract. A model is presented for the emergence of a primitive genetic code through the selection of a family of proteins capable of executing the code and catalyzing their own formation from polynucleotide templates. These proteins are assignment catalysts capable of modulating the rate of incorporation of different amino acids at the position of different codons. The starting point of the model is a polynucleotide based polypeptide construction process which maintains colinearity between template and product, but may not maintain a coded relationship between amino acids and codons. Among the primitive proteins made are assumed to be assignment catalysts characterized by structural and functional parameters which are used to formulate the production kinetics of these catalysts from available templates. Application of the model to the simple case of two letter codon and amino acid alphabets has been analyzed in detail. As the structural, functional, and kinetic parameters are varied, the dynamics undergoes many bifurcations, allowing an initially ambiguous system of catalysts to evolve to a coded, self-reproductive system. The proposed selective pressure of this evolution is the efficiency of utilization of monomers and energy. The model also simulates the qualitative features of suppression, in which a deleterious mutation is partly corrected by the introduction of translational error.

1. Introduction

Of the many complex problems associated with the origin of life, the origin of the genetic code is one of the most puzzling. The apparent universality of the code can be viewed as a consequence of a common ancestry of terrestrial life, while the remarkable accuracy of individual code assignments (i.e., codon-amino acid associations) can be ascribed to the specificity of aminoacyl-tRNA synthetases and other components of the translation system, resulting from the three-dimensional conformation of these macromolecules. In contrast, it is difficult to find a physical basis for the fact that the action of all aminoacyl-tRNA synthetases defines a coded relationship between the codon and amino acid alphabets. This relationship is the genetic code, which has the striking and important property of being unambiguous: only one amino acid is assigned to each codon, although several different codons may be related to the same amino acid (degeneracy).

The earliest attempts to understand the origin of this coded relationship were mechanistic theories, postulating stereochemical fits between amino acids and their codons or anticodons (Pelc and Welton, 1966). Although some specificity of interaction has been observed (Wagner and Arav, 1968; Saxinger and Ponnampuruma, 1971) the experimental evidence available to date does not substantiate a purely stereochemical theory of origin,

* Present Address: Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, U.S.A.

Copyright © 1982 by D. Reidel Publishing Co., Dordrecht, Holland, and Boston, U.S.A.
according to which a unique and universal code would be prescribed by interaction energies of amino acids and nucleotides. A purely stochastic theory of origin of the present-day code, postulating equiprobability of all codon–amino acid assignments and therefore all possible codes, may be discounted because of very low probabilities for the nucleation of a code for 20 amino acids through a 'frozen accident' (Crick, 1968; Eigen, 1971; Hoffman, 1975). This problem may be circumvented by assuming that code nucleation is required only for a few amino acids (or types of amino acids), further additions (or distinctions) being introduced through evolution. While these two theories are useful as extreme limiting cases, it is more likely that both stereochemical and stochastic effects, as well as special initial and boundary conditions, played important roles during the complex stages of molecular evolution leading to the nucleation of a primitive code.

An alternative approach is to emphasize the contribution of primitive proteins capable of modulating rates of amino acid-codon associations (assignment catalysts), and assume that the selective advantages of prebiotic systems derived from their kinetic features. This paper is an attempt to model the enzymatic hypothesis of code origin, which was suggested through the error catastrophe models of Orgel (1963, 1970), Hoffman (1974), Kirkwood and Holliday (1975), and Goel and Ycas (1975, 1976). The hypothesis focuses on the role of assignment catalysts, which correspond best to present day aminoacyl-tRNA synthetases, elongation factors, and ribosomal proteins, in establishing and maintaining coded translation as an essential process of living organisms. These models of error propagation have already shown, that within the context of a single code, there is a threshold of accuracy above which the system can achieve stable and correct translation, but unfavorable initial conditions can drive the system to a stable but error laden state.

Mizutani and Ponnampemuma (1977) have already presented an enzymatic model for the evolution of a translation machinery in which polymerases are characterized by activities and selectivities (expressed as the grade of a polymerase) and a number of critical sites determining: these properties. Each polymerase is assumed to belong to a genetic code, and transition probabilities between different grade systems are defined. Their simulation results indicate that for reasonable values of selectivities and transition probabilities, the highest grade system will eventually dominate. However, the model treats transition probabilities between different grade components as arbitrary parameters, and does not reflect their dependence on the existing population of polymerases. Furthermore, it seems that transitions between systems that belong to the same or alternate codes need to be distinguished.

Our purpose is to construct a kinetic description of a system of assignment catalysts, such that assignment probabilities are expressed in terms of concentrations and activities of existing catalysts, and all possible codes (for a given choice of amino acid and codon alphabets) are explicitly accounted for. In accordance with the many analyses of the patterns of degeneracy in the code (Crick, 1968; Jukes, 1973, 1974; Walker, 1974; Wong, 1975, 1976), we will assume that the primitive code recognized (or distinguished) only a few (2–6) amino acids and codons.