ISSN 2075-1729
www.mdpi.com/journal/life

Peer-Review Record:

The Hypothesis that the Genetic Code Originated in Coupled Synthesis of Proteins and the Evolutionary Predecessors of Nucleic Acids in Primitive Cells

Brian R. Francis

***Life* 2015, *5*, 467-505, doi:10.3390/life5010467**

Reviewer 1: Anonymous
Reviewer 2: Anonymous
Editor: Pabulo H. Rampelotto (Editor in Chief of Life), Hyman Hartman and Temple F. Smith (Guest editor of Special Issue "The Origin and Evolution of the Genetic Code")

Received: 8 December 2014 / Accepted: 2 February 2015 / Published: 11 February 2015

First Round of Evaluation

Round 1: Reviewer 1 Report and Author Response

This represents an intriguing hypothesis for the origin of the genetic code and the co-evolution of proteins & nucleotides. The author takes on the evolution of energy transduction, polymer replication, reaction catalysis and most important getting that first membrane around the stuff of life. The author leads us through central metabolic pathways and makes the interesting postulation of using thioester bonding in polymer linkage. The discussion of each phase is thorough. I found the work fascinating and intellectually stimulating to the point where I can't find fault with the persuasive argumants. Unfortunatly Figure 2 downloaded as gibberish, but I can get the jist from the legend. It was a fascinating read, well-presented.

Response: Thank you for your support.

Round 1: Reviewer 2 Report and Author Response

This paper provides a quite interesting hypothesis, suggesting that the earliest predecessor of the nucleic acids was a β-linked polyester made from malic acid. The author considers that the side chains in the β-linked polyester are carboxylic acid groups capable of forming interstrand double hydrogen bonds and evolution of the nucleic acids involved changes to the backbone and side chain of poly(β-D-malic acid). In addition, the author provides an idea suggesting that the singlet coding system for replication evolved into a four nucleotide/four amino acid process (AMP = aspartic acid, GMP = glycine, UMP =

valine, CMP = alanine) and then into the triplet ribosomal process that permitted multiple copies of protein to be synthesized independent of replication, as the synthesis of nucleic acids evolved from β-linked polyesters.

But, I would like to recommend the author to add your ideas and/or opinions for my comments as described below, if possible.

Response: Thank you for comments. My responses to the questions raised are given below.

(1) I think that it would be difficult to form stable double stranded structure with the β-linked polyester chain, since all bonds on the main chain of the polymer are single-bonds which rotate freely around the bonds. How do you reply about the difficulty?

Response: The Reviewer asks an important question to which I have been unable to find an answer. A β-polyester strand with carboxylic acid side chains would have entropy due to freedom of rotation around its backbone bonds that would be lost if it were to become part of a double strand in which double hydrogen bonds are formed between the carboxylic acid side chains. The Reviewer is asking whether the enthalpy gained from formation of double hydrogen bonds in a double strand would offset the loss of entropy from the single stranded polymers. Poly(β-L-malic acid) is a biosynthetic product. Although poly(β-L-malic acid) has been studied in some detail, to my knowledge the presence or absence of double stranded interactions has not been investigated. I have added a discussion of this issue to section 17.

(2) How and which kinds of genetic information could be written on the β-linked polyester chain? I think that any meaningful genetic information could not be obtained on the polymer.

Response: Genetic information was initially carried in the different side chains of the β-polyester as shown in Figure 5A,B.

(3) Random copolymer of glycine and alanine would not be folded into water-soluble globular structures, which is pre-requisitely required to form catalytic activity on a protein surface, because hydrophobic amino acid as valine is not contained in the polymer.

Response: One of the main features of this hypothesis is that coupled synthesis of proteins and nucleic acids could have produced coded rather than random proteins. Whether or not proteins were coded, the Reviewer is correct that a protein comprised of Gly and Ala would not form globular structures. I think it is more likely that coded Gly-Ala proteins were inserted into the membrane material (see Section 19). Gly-rich loops, perhaps near the N-terminal and C-terminal ends of proteins and close to the inner surface of the membrane, could have formed binding sites for catalysis (see Francis BR. 2013, J Mol Evol 77, 134–158; Milner-White EJ, Russell MJ. 2011, Genes 2, 671–678; Bernhardt HS, Patrick,WM. 2014 J Mol Evol 78, 307–309).

(4) Any genetic information for formation of amino acid sequence should be destroyed during transfer from poly(β-D-malamide) to RNA backbone as drawn in Figure 4. How do you explain about that?

Response: Figure 4 considers changes in the backbone separately from changes in the side chain, which are discussed in relation to Figures 5A,B and 6. Information is not destroyed during evolution of the *backbone. Because this point may not have been sufficiently clear, I have added a sentence to the legend for Figure 4.*

(5) Any genetic information also should be lost during evolution of the singlet coding system into the triplet coding system. How do you explain about that?

Response: In my 2011 paper in Trends in Evolutionary Biology, which is referenced in this paper, I explain how continuity of coding could have been maintained during the transition from the singlet coding system to the triplet coding system. I refer the Reviewer to this paper for details.

(6) The idea supposing that the genetic code would originated from that encoding four amino acids, Gly, Ala, Asp and Val, has already proposed by Ikehara *et al.* (*J. Mol. Evol*., Vol. 54, 530–538 (2002), Higgs (Biol Direct. 2009 Apr. 24;4:16.) and van der Gulik *et al.* (*J Theor Biol*. 2009 Vol. 261, No. 4:531–539). So, the author should cite the papers above at Section 23.

Response: I apologize for not properly referencing these authors and have done so in the revision. At the same time I have also added references to the paper by Eigen and Winkler-Oswatitsch (1981) that first proposed that early proteins were composed of Gly, Ala, Val and Asp, and my own 2000 paper that incorporated their proposal.