July 16-21, 2017 at UC San Diego, CA, USA

A proposal of the Ur-proteome

Miryam Palacios-Pérez[a], Fernando Andrade-Díaz[a], Marco V. José[a,*]
a *Universidad Nacional Autónoma de México, Instituto de Investigaciones Biomédicas, Ciudad de México, D.F., México*

* *Corresponding author*, marcojose@biomedicas.unam.mx

Abstract:

We can uncover the plausible Ur-proteome encoded in the RNY chains (where R indicates purine, N means any of the four bases, and Y indicates pyrimidine), that Eigen and Schuster suggested as the first genetic code in the early 70's, based on the logical deduction of the thermodynamic properties of that kind of polyribonucleotide existing in the RNA world.

The evolution of that primeval code has been previously investigated under rigorous mathematical approaches, and it was demonstrated how the current standard genetic code (SGC) can be derived via a dichotomous pathway starting from the RNY code.

Presently, we could find vestiges of the primeval phenotype, corresponding to the RNA genotype, as a collection of peptides constructed from the 8 amino acids (aa) encoded in the 16 RNY triplets. We began by extracting an RNY genome from a contemporary organism and the resulting smaller genome was used as a query in order to obtain a list of ancient proteins encoded by RNY codons, which at first instance was mysteriously heterogeneous.

By looking at the fragments encoded by RNY triplets, it was noteworthy that they are positioned, not in catalytic sites, but in the cofactor binding sites. Some fragments were then extracted, their three-dimensional structure was predicted and, without any additional manipulation, it was startling that such peptides actually bind somehow the now called cofactors, which has been proposed among the early prebiotic molecules nonetheless. It is necessary to recall that currently these fragments contain not only the 8 aa of the ancestral phenotype but they are composed mainly of them in diverse combinations.

Notwithstanding our methodology does not fit into the classic “bottom-up” neither “top-down” approaches, our approach that uses only the 16 RNY triplets as genotype and as phenotype the corresponding 8 aa, an ***Ur-proteome*** can be wrought that consists of a set of primordial peptides that work as ***Cofactor Stabilising Binding Sites (CSBSs)***, i.e. the primitive ***bindome***. It implies that the stabilization of a molecule appeared long before its catalytic use. Indeed, such notion of CSBSs as the first proteins modules in progenotes is not unreliable, and is congruous with several propositions about the primitive forms of life.

Finally, we can state that the ancestral CSBSs constitute the primordia of the peptides that would eventually evolved –likely aided by HGT (thus not purely by Darwinian processes)–, yielding the basic repertoire of LUCA.