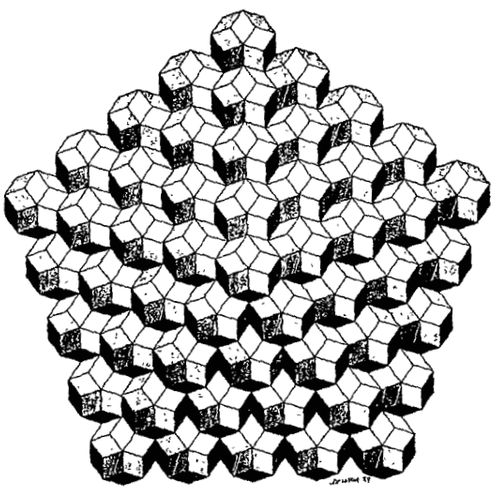
Symmetry of STRUCTURE

an interdisciplinary Symposium

Abstracts

II.



Edited by Gy. Darvas and D. Nagy

Budapest

August 13-19, 1989

hungary

THE CRITERION OF SIMMETRY-ASYMMETRY IN THE EVOLUTION OF THE GENETIC CODES

C. PORTELLI, A. PORTELLI

Department of Biophysics, Faculty of Medicine, 8, P.Groza Blvd, 76241Bucharest, Romania

According to the criterion of symmetry-asymmetry, the triplets of the genetic code can be divided in four classes: 1) In the first class there are the codons which present in the first and third positions the same nucleotide base. For instance, AGA and GAG; 2) In the second class there are codons which present in the first and third position two different nucleotide bases, but both purines, or both pyrimidines. For instance, AGG and GAA. 3) In the third class there are codons which present in the first and third positions two complementary nucleotide bases. For instance, AGT and CGG. 4) In the fourth class there are codons which present in the first and third positions two anti-complementary nucleotide bases. For instance, GAT and ATC.

The criterion of symmetry-asymmetry is very important for the classification of codons. The following arguments plead in this sense: 1) In each class there are 16 codons. The codons of the second class can be deduced from these of the first class by the substitution of the third nucleotide base according to the rule: A $\longrightarrow$ G, G $\longrightarrow$ A, C $\longrightarrow$ U and U $\longrightarrow$ C.

The codons of the fourth class can be deduced from these of the third class by means of a similar rule to that utilized for the transition from the first to the second class.

2) If we make abstraction from the differences existing between the two purine bases, or these between the pyrimidine bases, the codons of the first and second classes are symmetrical codons, while these of the third and fourth classes are symmetrical triplets. There are 16 pairs of symmetrical codons and 16 pairs of asymmetrical codons. The codons of each pair specify the same amino acid.

In the literature there are mentioned three types of genetic codes. The "universal" code, which is utilized by all living organisms, starting from viruses and ending with man. The mammalian mitochondrial code and yeast mitochondrial code, which are utilized by the respective mitochondria (Tzagoloff, 1982; Grivell, 1983).

From the point of view of symmetry, the code of mammalian mitochondria shows the most harmonious structure, the code of yeast mitochondria is in an intermediate position, while the "universal" code presents the most asymmetrical form.

On the other hand, the molecular mechanisms involved in the processes of translation and transcription present the most simple form in the genomes which utilize the mammalian mitochondrial code and the most complex form in the genomes which utilize the "universal" genetic code (Grivell, 1983).

Taking this into account the code of mammalian mitochondria is probably to preserve the most related structure to the primitive genetic code, while the "universal" genetic code shows the most evolved form. Therefore, the evolution of the genetic codes developed from a more symmetrical form to an asymmetrical structure.

In the code of mammalian mitochondria, in each class of codons there are: a) one codon which initiates the translation; b) one codon which ends the translation; c) different codons for hydrophobic and hydrophilic amino acids.

Therefore, each class of codons of mammalian mitochondrial code could function as an independent genetic code. This fact suggests the hypothesis that during the evolution of the genetic codes a phenomenon of duplication of codons took place. The 16 codons of the first class have generated the 16 codons of the second class. Then, the 16 pairs of symmetrical codons have generated by a new duplication the 16 pairs of asymmetrical codons.

In the genetic codes there is a structural and functional correspondence between the second nucleotide base of the codons and the specific amino acids. If the second nucleotide base is a purine, then the specified amino acid is hydrophilic. If the second nucleotide base is a pyrimidine, the specified amino acid is hydrophobic.

Taking into account these aspects, the following model concerning the origin of the genetic codes can be conceived.

At the origin of life, certain primitive polynucleotides have been achieved by means of some non-biological ways. These primitive polynucleotides have been formed by ribonucleotides, because only the ribonucleotides can be synthetized without the intervention of some specific enzymes (Eigen, 1977).

Despite of the fact that they have been ribonucleotides, these primitive polynucleotides contained thymine instead of uracyl. This substitution was necessary because thymine is more suitable to achieve a direct hydrophobic interaction with the hydrophobic amino acids, then uracyl. For the existence of these primary polynucleotides at the origin of life pleads the fact they are still present like vestiges in the sequences of the actual mammalian mitochondrial genes (Grivell, 1983).

In the structure of the primary polynucleotides there were three principal zones: a) an initial sequence formed by pyrimidine bases; b) an intermediary sequence formed by an alternate succession of purine-pyrimidine bases; c) a final zone formed by purine bases.

The initial zone was necessary for the initiation of synthesis, because the hydrophobic interactions are more suitable to achieve intermolecular connections with hydrophobic amino acids, than the hydrogen bonds with the hydrophilic amino acids.

The intermediate zone was necessary because the polynucleotides having the type PpPpPpP (purine-pyrimidine-purine) can perform easely the transition from a conformation to another one (Felsenfeld, 1985).

The final zone was necessary because it determined the synthesis of a hydrophilic tail at the end of the polypeptide, which was important for the building of different biological membranes.

The codons existing in these three zones have been symmetrical codons (PPP, PpP, pPp and PPP).

According to the data presented in literature, it is probably that the purine bases have selectively connected the hydrophilic amino acids, while pyrimidine bases have fixed the hydrophobic amino acids.

After the fixation of amino acids at the level of the nucleotide bases, some primitive molecular mechanisms have acted and have produced the synthesis of the respective polypeptide. Then, the primary polypeptide has acted on the respective polynucleotide and has induced its replication. In this way, the first structural and functional correspondence between the primitive polynucleotides and polypeptides has been achieved.

The role of the triplets was determined by the fact that the selectivity of a nucleotide base, for certain amino acid was also influenced by the neighbouring situated nucleotide bases (the previous and the next in sequence).

For instance, if cytosine was included between two thymine bases, it had a smaller probability to connect a hydrophobic amino acid, because the thymine bases present a greater affinity for the hydrophobic amino acid than cytosine. Due to this, in the actual genetic codes the triplet TCT specifies a hydrophilic amino acid (serine).

In conclusion, at the origin of the genetic codes an important role had been played by the direct interaction between the nucleotide bases and the amino acids, and by the symmetrical codons.

REFERENCES

Eigen M., P. Schuster, (1977), Naturwissenschaften, 64, 541.

Fox S.W., K. Doese, (1977), Molecular evolution and the origin of life, Plenum Press, New York.

Grivell L., (1983), Scient. Am., 246, 60.

Lehninger A., (1982), Biochemistry, Werth Publishers, New York.

Portelli C., A. Portelli, Daniela Cigani, (1985), Rev.roum.Physiol. 23, 139.

Tzagoloff A., (1982), Mitochondria, Plenum Press, New York.