life

MDPI

Article

Emergence of a "Cyclosome" in a Primitive Network Capable of Building "Infinite" Proteins †

Jacques Demongeot [1,*] and Vic Norris [2]

[1] Faculty of Medicine, Université Grenoble Alpes, AGEIS EA 7407 Tools for e-Gnosis Medical, 38700 La Tronche, France

[2] Laboratory of Microbiology Signals and Microenvironment, Université de Rouen, 76821 Mont-Saint-Aignan CEDEX, France; victor.norris@univ-rouen.fr

* Correspondence: jacques.demongeot@univ-grenoble-alpes.fr; Tel.: +33-476-637-135

† The present paper is dedicated to René Thomas (1928–2017) who was the first to encourage in France–especially via the Selection Committee of the Institut Universitaire de France—this type of research on RNAs and the origin of life. The research undertaken by René Thomas himself ranged widely from RNA biochemistry (notably the denaturation/renaturation of nucleic acids) to genetics (in particular that of phages), theoretical biology, and system dynamics. René Thomas was an extraordinarily creative scientist and an unwavering friend with his close colleagues, giving precious advice on all aspects of their scientific and ordinary lives.

Received: 1 May 2019; Accepted: 13 June 2019; Published: 18 June 2019

Abstract: We argue for the existence of an RNA sequence, called the AL (for ALpha) sequence, which may have played a role at the origin of life; this role entailed the AL sequence helping generate the first peptide assemblies via a primitive network. These peptide assemblies included "infinite" proteins. The AL sequence was constructed on an economy principle as the smallest RNA ring having one representative of each codon's synonymy class and capable of adopting a non-functional but nevertheless evolutionarily stable hairpin form that resisted denaturation due to environmental changes in pH, hydration, temperature, etc. Long subsequences from the AL ring resemble sequences from tRNAs and 5S rRNAs of numerous species like the proteobacterium, *Rhodobacter sphaeroides*. Pentameric subsequences from the AL are present more frequently than expected in current genomes, in particular, in genes encoding some of the proteins associated with ribosomes like tRNA synthetases. Such relics may help explain the existence of universal sequences like exon/intron frontier regions, Shine-Dalgarno sequence (present in bacterial and archaeal mRNAs), CRISPR and mitochondrial loop sequences.

Keywords: primitive network; cyclosome; stereochemical hypothesis; small acid-soluble proteins; tRNA synthetases

1. Introduction

After the first observations of a ribosome sixty years ago by G.E. Palade and P. Siekevitz [1], theoreticians proposed the first models about the origin of life using tools from statistical mechanics: In 1967, S. Ulam simulated large automata networks and remarked that, with simple growth rules, he obtained complicated patterns similar to those observed in biology [2]. In 1968, inspired by these results, J. Conway started to do similar simulations of cellular automata, in particular a new one called the "Game of Life" by M. Gardner in 1970, because it showed discrete numerical structures moving on a plane and being duplicated [3]. Independently, in parallel, a French team composed of two biologists and one physician-mathematician (J. Besson, P. Gavaudan, and M.P. Schützenberger) worked on the optimality of the genetic code and the existence of primitive RNAs [4] similar to the invariant parts of tRNA loops (see supplementary material in [5]). Unfortunately, Conway's algorithm

did not incorporate realistic genetic considerations in the game that would have made it possible to test, for example, the plausibility of molecular events hypothesized to have led to the appearance of life. In the spirit of such games, we revisit all the above works a half century later by proposing simple RNA structures that could have served as matrices for building first peptides and that constituted what we term here "a cyclosome".

Since 1984, many papers, including in particular those of Yonath and collaborators, have explored the possibility of a chain of ancestors of the present ribosomal molecules and have proposed for these primitive structures the name of "protoribosome" [6–23]. More recent papers have similarly revisited the evolution of the tRNA and aminoacid-tRNA synthetases (or ligases) in the ribosomal ecosystem [24,25]. In the present paper, we try to complement these structural approaches by a pure informatics approach showing that from simple realistic constraints, sequences and secondary RNA structures can be found with selectable properties of resistance to denaturation and of storage of the genetic code. We start the corresponding "game of life" in Section 2 by searching for a realistic primitive genetic network regulating the dynamics of the first actors involved in making peptides. First, we designate an RNA ring, termed AL for ALpha, as the central actor, and then we search for AL candidates in rings having a minimal length and possessing one and only one codon from each synonymy class of the genetic code; these requirements are chosen in order to favor a stereochemical neighborhood of the AL rich in amino acids, and hence favor the synthesis of an 'infinite' (theoretically endless but randomly cleaved by environmental factors) protein that would have allowed evolution to select a great many peptides and proteins from the making and subsequent breaking of peptide bonds with a degree of randomness. In order to ensure the survival of the AL in the absence of amino acids and hence in the absence of a selectable function, we ask more of the AL, namely, that it should be able to adopt a hairpin structure that, optimally, would be as stable as possible, have the smallest head (three nucleotides) and the longest tail (9 pairs of nucleotides). In Section 3, we search amongst those species that have two or more types of short RNA sequences (e.g., 5S and 29S ribosomal RNA and the loops of transfer RNA) for those that share with the AL at least one nonameric (9-mer) sequence consistent with their co-evolution. We report in Section 3 those species that satisfy this constraint, e.g., a proteobacterium, *Rhodobacter sphaeroides* (Figure 1A,B). By combining these sequences, we construct a circular AL ring with 22 nucleotides and then search for the most stable hairpin with the same sequence as the AL in large genetic data bases with repeated motifs from the AL, namely pentameric subsequences. Then, in Section 4, we look for AL relics in current Archaea genomes and in some ancient structures like ribozymes. Finally, in Section 5, we propose how the AL arose from catalysis by interfaces between membrane domains and how the AL may have generated "infinite" proteins as part of its role in the evolution towards a "protein-synthesizing machine" in its own right (perhaps 'the ancestral protein-synthesizing machine') that we term a "cyclosome". The latter was facilitated by the production of the first nucleo-peptide conjugates as shown by the frequency of the pentameric relics of the AL which serves as a scalar for proximity to AL.

2. A Primitive Network at the Origin of Life

In our hypothesis, amino acids were concentrated around the AL, which acted as a "proto-nucleus" to allow the first "organ" or "cyclosome" to synthesize peptides. Insofar as an object corresponds to a discontinuity in a field of connectivity [26], the boundary of this cyclosome corresponded to a discontinuity in the gradient of peptides around the AL.

The boundary of the first functional "machine" able to build peptides can be defined as a peptide gradient boundary centred on the "proto-nucleus" AL, resulting from an amino acid confinement around the AL favoring the occurrence of peptide bonds. This "organ" functioned as a "cyclosome" in a "proto-membrane", thus as a "proto-cell" with a circular organization. This proto-cell is a solution to the problem of how to obtain autopoiesis: Peptide synthesis favored by the AL was necessary to repair the proto-cell membrane made of hydrophobic peptides and lipids, which reciprocally protected the AL against denaturation by ensuring the integrity of the proto-nucleus. The autopoiesis network

underlying this organization has been studied in [27–29] and exhibits exponential growth if the peptide proto-membrane allows the entry of nucleic acids for AL replication. We can represent its dynamics by defining the variables of the network and their interactions using a system of differential equations (1) whose Jacobian graph is given in Figure 1: Let us denote by R, A, B, E, M, and P for, respectively, the concentration of AL Ring, Amino acids, nucleotide Bases, hydrophilic Enzyme peptides, hydrophobic Membrane peptides, and the Pool of lipids plus the elements C, N, and H_2O:

$$\begin{aligned} dR/dt &= d_R\Delta R + k_B B - k_R R \\ dA/dt &= d_A\Delta A + k_P P - k_A RA - k'_E RA \\ dB/dt &= d_B\Delta B + k'_P P - k_B B \\ dE/dt &= d_E\Delta E + k'_E RA - k_E E \\ dM/dt &= d_M\Delta M + k_A RA - k_M M \\ dP/dt &= d_P\Delta P + k_R R + k_M M - k'_P P - k_P P \end{aligned} \quad (1)$$

In the absence of diffusion (diffusion coefficients d_i's equal to 0), the differential system (1) has an initial exponential growth behaviour and tends, if P is constant and initial values of variables are not zero (which corresponds to an unstable steady state), towards the unique stable stationary state:

$$(R^*,A^*,B^*,E^*,M^*) = (K/k_R,\ K'k_R/K(k_A + k'_E),\ K/k_B,\ K'k'_E/k_E(k_A + k'_E),\ K'k_A/k_M(k_A + k'_E)) \quad (2)$$

with the following Jacobian matrix J* equal to:

$-k_R$	0	k_B	0	0
$-K'k_R/K$	$-K(k_A+k'_E)/k_E$	0	0	0
0	0	$-k_B$	0	0
$K'k_Rk'_E/K(k_A+k'_E)$	Kk'_E/k_R	0	$-k_E$	
$K'k_Rk_A/K(k_A+k'_E)$	Kk_A/k_R	0	0	$-k_M$

whose characteristic polynomial $(k_R + \lambda)(K(k_A + k'_E)/k_E + \lambda)(k_B + \lambda)(k_E + \lambda)(k_M + \lambda) = 0$ has only negative eigenvalues, ensuring the stability of the stationary state (2) (R*,A*,B*,E*,M*). If we add a diffusion term for the different metabolites, this dynamics leads to the spatial segregation of R, A, B, M, and P into structures like a protein-synthesizing machine made of the AL or the anti-sense AL (R), proto-cytoplasm (A, B), proto-membrane (M), and building blocks (P). The AL serves as a template for the formation of hydrophilic enzymatic peptides E (Figure 1) able to activate (or inhibit depending on their catalytic properties) the AL or anti-sense AL replication with a reaction constant k_r [30]. If we introduce diffusion processes (whose viscosity coefficients depend on the membrane concentration M) in the purely reaction differential system, we get the diffusion-reaction equations (1), for which a close discrete analogue has been already simulated in [31,32], and which shows a progressive space segmentation by the M gradient.

During its exponential growth and diffusion, the boundary of the system (1) is chosen as the gradient boundary of peptides polymerized from amino acids. Growth stops in the case of a lack of nucleic acid or protein precursors, i.e., because of the disappearance of the elements of the C, N, and H_2O pool (which provides the amino acids and nucleotides that are consumed during the growth).

3. Construction of the AL RNA Ring

We have shown by using constraint programming and a step-by-step computation [33,34] that only 25 RNA rings satisfy the following constraints:

- All dinucleotides should appear at least once (apart from CG because of CG suppression).
- Among rings satisfying the constraint "to be as short as possible and contain at least one codon of each amino acid synonymy class", there is no solution for a length below 22 nucleotides. For length 22, 29,520 solutions contain the codon AUN twice, N being G for 52% of the solutions.
- From the 29,520 solutions, only 25 rings allow the formation of a hairpin at least 9-bases long.

- Of these 25 rings, 19 have both start and stop codons.
- Through calculation of the average genetic distances to the others (e.g., circular Hamming distance, permutation distance, and edit distance), one singular ring exhibits a minimum distance as compared to the others. Only one sequence, called AL (for ALpha) is thus acting as the barycenter of the set of the 18 others: 5′-AUGGUACUGCCAUUCAAGAUGA-3′.

Then, we remark that AL appears by merging the following sequences of the genome of *Rhodobacter sphaeroides* (Figure 2): AATGGTACTTCCATTCGATATG from the Gly-tRNA^{TCC} loops, AATGGTACTGCGTCTCAAGACG from 5S rRNA [35].

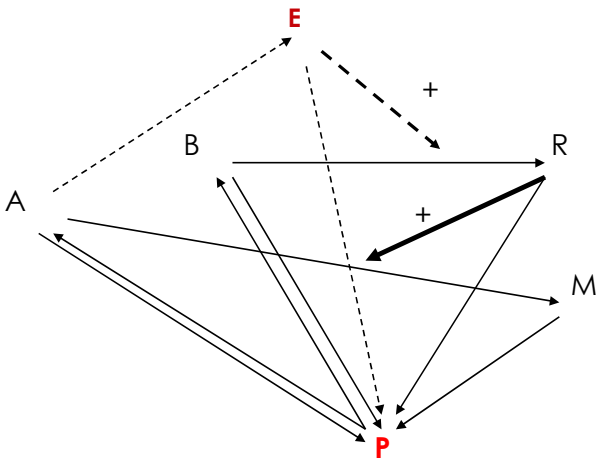

Figure 1. Interaction graph of the genetic network of an autopoiesis model inspired from P. Bourgine and J. Stewart [28] with only activation arrows except the dashed arrow, which can represent either an activation or an inhibition. P (in red) represents the Pool of the elements C, Nand H_2O, E (in brown) hydrophilic Enzyme peptides, R AL Ring, (**A**) Amino acids, (**B**) nucleotide Bases, and M hydrophobic Membrane peptides.

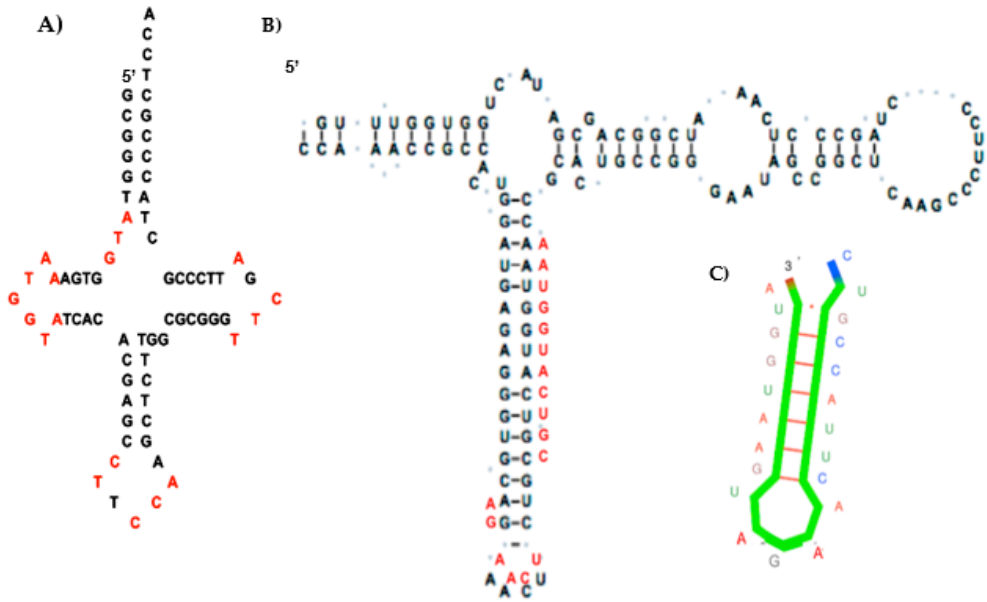

Figure 2. (**A**) AL subsequences (in red) ATG, AATGGTA, CT, and CCATTC from the loops of the Gly-tRNA^{TCC} of *Rhodobacter sphaeroides*; (**B**) AL hemi-sequence AAUGGUACUGC (in red) and AL-hexamer UCAAGA (in red) from the hairpin of the 5S rRNA of *Rhodobacter sphaeroides* (adapted from [35]); (**C**) Optimal hairpin form for AL (from Kinefold [36]).

It is possible to design, by using the Kinefold® algorithm [36], the most thermodynamically stable hairpin (Gibbs free energy equal to $\Delta G = -9.5$ kcal/mol in Figure 2) among the 22 RNA chains obtained from the circular permutations of AL (Figure 2C). This structure could explain why, during denaturation, there is first a loss of the AL-hexamer CUGCCA (anticodon loop of current Gly-tRNA^{GCC}s) and then a break between AL-heptamers UUCAAGA (the TΨ-loop of current tRNAs) and AAUGGUA (the D-loop of current tRNAs). An argument in favor of this scenario is the distribution of the pentamer frequencies inside the current genome (from Rfam database, http://rfam.xfam.org/), which shows the two highest survival probabilities for the AL-pentamers coming from the most stable part of AL, also parts of the D-loop and TΨ-loop of the present tRNAS, i.e., AAUGG, AUGGU, UGGUA, GGUAC, TTCAA, TCAAG, and CAAGA. If we consider other subsequences of AL, we find many repeated motifs, such as AATGG [37] and GATG [38] from human microsatellites, AGAT from vertebrate repeated UTR motifs [39], and CCATTCA from the Alpha Satellite of Human Chromosome 17 [40] and from the HMG box (High Mobility Group Box, a protein domain involved in DNA binding [41]), as well as the optimal codons that determine mRNA stability in the yeast genome [42].

We can generalize the result obtained from *R. sphaeroides* to other archaeal, bacterial, and eukaryotic species as shown in Table 1. The genetic code consists of 64 triplets made of 3 letters representing purine bases—A for Adenine and G for Guanine—and pyrimidine ones—U for Uracil and C for Cytosine—that can be grouped into 21 synonymy classes.

Table 1. Common subsequences among genomes of different species.

Rhodobacter sphaeroides
AATGGTATTCCCATTCGATTTG tRNA-Gly (http://lowelab.ucsc.edu/GtRNAdb/Rhod_spha_ATCC_17029/rhodSpha_ATCC17029-tRNAs.fa)
AATGGTACTGCGTC**TCAAGACG 5S rRNA (http://www.combio.pl/rrna/alignment/)**
CC**TGG**A**ACTGCCATT**G**AA**ACT**C 16S rRNA (https://www.ncbi.nlm.nih.gov/nuccore/636559472?report=fasta)**
AATGGTACTGCCATTCAAGATG Consensus
Rhodospirillum rubrum
TGAATGGTACTTCCAATTCGA**A tRNA-Gly (http://trna.ie.niigata-u.ac.jp/)**
CC**AATGGTACTGC**GTC**TTAAGG 5S rRNA (http://www.combio.pl/rrna/alignment/)**
CTCCA**GGTACTGCCCTTGA**TAC **16S rRNA (https://www.arb-silva.de/browser/)**
C**GAATGGTACTGCCATTTAAAA Consensus**
Rubellimicrobium thermophilum
A**GTGGTACTTCCATTCGACATG tRNA-Gly (http://trna.ie.niigata-u.ac.jp/)**
AATGGTACTGCGCC**TCAAGACG 5S rRNA (http://www.combio.pl/rrna/alignment/)**
G**ATGGTCCAGGCGCTGCCGCTC 16S rRNA (https://www.arb-silva.de/browser/)**
AATGGTACTGCCACT**CAAGATG Consensus**
Haematobacter missouriensis
A**GGGGTATTGCCATTCGAATTA tRNA-Gly (http://trna.ie.niigata-u.ac.jp/cgi-bin/trnadb/whole_detail.cgi?SID=2138813)**
T**ATGGTGCTTCCATTCCCGCTA tRNA-Gly (https://www.ncbi.nlm.nih.gov/nuccore/672903602?report=genbank)**
AATGGTACTGCGTC**TCAAGACG 5S rRNA (http://www.combio.pl/rrna/alignment/)**
AATGGTAGTGACAATGGGTTAA 16S rRNA (https://www.arb-silva.de/browser/)
AATGGTACTGCCATTCAAGATG Consensus

Table 1. *Cont.*

Paracoccus sp. S4493
AATGGTACTTCCCTTCGATTTA tRNA-Gly (https://www.ncbi.nlm.nih.gov/nuccore/NZ_JXYF01000001.1?from=63131&to=63204&sat=19&sat_key=63645080&report=fasta&strand=2)
GATGGTACTGCGTCTTAAGACG 5S rRNA (http://www.combio.pl/rrna/alignment/)
AATGGTGGTGACAGTGGGTTAA 16S rRNA (http://www.ebi.ac.uk/ena/data/view/FJ457300&display=fasta)
AATGGTACTGCCATTCAATTTA Consensus
Flavobacteria bacterium MS024-2A
CTGGTATTGCCATTCGAATCGC tRNA-Gly (http://gtrnadb.ucsc.edu/genomes/bacteria/Flav_bact_3519_10/flavBact_3519_10-tRNAs.fa)
ATGGTACTGCCATCCGGTGGGA 5S rRNA (http://www.combio.pl/rrna/alignment/)
ATGGTAACGGCATACCAAGGCA 16S RRNA (http://www.ebi.ac.uk/ena/data/view/AM931128&display=fasta)
ATGGTACTGCCATTCGAAGGGA Consensus
Methanococcus maripaludis
CTGGTACTTCCATTCAAATCGT tRNA-Gly (http://gtrnadb.ucsc.edu/genomes/archaea/Meth_mari_C5/methMari_C5_1-tRNAs.fa)
TAAGTACTGCCATCUGGUGGGA 5S rRNA (http://biobases.ibch.poznan.pl/htbins/getseq.cgi?name=Methanococcus%20maripaludis)
TCGGTACGGGCCTTGAGAGAGG 16S rRNA (http://www.ebi.ac.uk/ena/data/view/AB546258&display=fasta)
TTGGTACTGCCATTCAGAGAGA Consensus
Tremella mesenterica
GATCTGCGAAGTCAAGATGAAT 5S rRNA (http://www.combio.pl/rrna/alignment/)
GGTAATTCTAGAGCTAATACAT18S rRNA (https://www.ncbi.nlm.nih.gov/nuccore/256600119?report=fasta)
GTACCGTGAGGGAAAGATGAAA 28S rRNA (https://www.ncbi.nlm.nih.gov/nuccore/46402656?report=fasta)
GGTCCGTGAAGTCAAGATGAAT Consensus
Homo sapiens
GTGGTACTCCCATTCAATTTGGtRNA (http://trna.bioinf.uni-leipzig.de/DataOutput/Result)
ATGGTAGTCGCCGTGCCTACCA 18S rRNA (https://www.ncbi.nlm.nih.gov/nuccore/225637497?report=fasta)
ATGGTAATCCTGCTCAGTACGA 28S rRNA (https://www.ncbi.nlm.nih.gov/nuccore/1154886866?report=fasta)
ATGGTACTCCCATTCAATACGA Consensus
AATGGTACTGCCATTTAAAACG Consensus Bacteria
AATGGTACTGCCATTCAAGATG Consensus Bacteria
AATGGTACTGCCATTCAAGATG Consensus Bacteria
AATGGTACTGCCACTCAAGATG Consensus Bacteria
AATGGTACTGCCATTCAAGATG Consensus Bacteria
AATGGTACTGCCATTCAATTTA Consensus Bacteria
ATTGGTACTGCCATTCAGAGAG Consensus Archaea
AATGGTCCGTGAAGTCAAGATG Consensus Eukaryote AATGGTACTCCCATTCAATACG Consensus Eukaryote
AATGGTACTGCCATTCAAGATG Consensus consensorum

Each class contains between 1 and 6 triplets; 20 classes correspond to the 20 amino acids (except for one class containing only 1 triplet, which corresponds either to the amino acid Methionine or, if this triplet initiates a sequence of messenger RNA (mRNA), to a "start" punctuation symbol), plus one class corresponding to the "end" punctuation symbol terminating the mRNA sequences. It has been shown that stereochemical bonds can favor a non-permanent, reversible link between amino acids (AA) and codons or anticodons of their AA synonymy class [43–47]. The 25 selected rings satisfy two opposite constraints corresponding to a min-max problem: (i) to be as short as possible, and (ii) to contain one and only one triplet corresponding to each amino acid synonymy class. The latter constraint would allow the rings to serve as a "matrimonial agency" concentrating amino acids in the vicinity of the ring and thereby favoring the links between any pair of them via peptide bonds [48–55]. The 25 RNA rings selected can be considered as ancestors of the tRNA of the 22 AAs including Pyrrolysine and Selenocysteine (Figure 3), with Serine counted twice, and Tyrosine and Aspartic Acid able to replace C by U in their tRNA anticodons [56,57].

The 12 rings in red in Figure 2 could correspond to an intermediary genetic code using the wobble mechanism present in Archaea [58–60] and many other organisms [61,62]. The AL ring (resp. AL' anti-ring) selects and confines more L-aminoacids (resp. D-aminoacids) and catalyses the synthesis of either hydrophobic or hydrophilic peptides [63–65]. We can note that in [9] peptide synthesis was achieved experimentally by using as RNA template a heptameric subsequences of AL, AAUGGU.

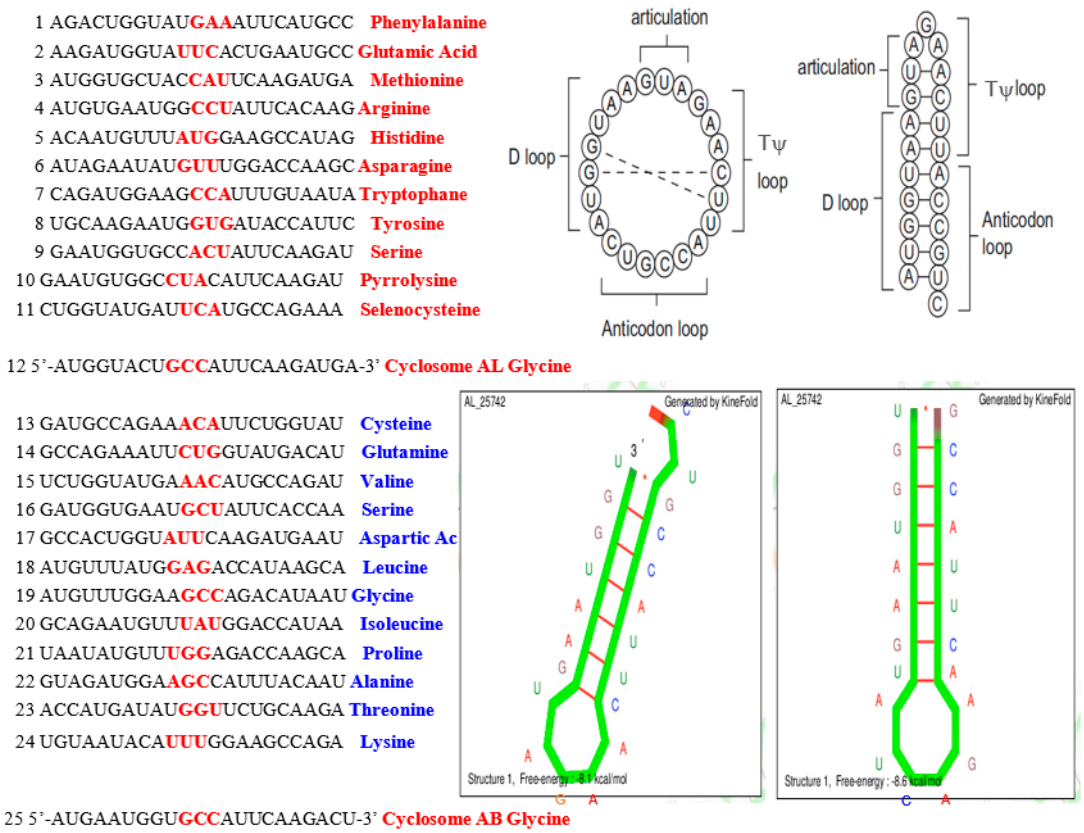

Figure 3. Left: The 25 RNA sequence candidates for the archetypal tRNAs that bound the 22 amino acids. Top Right: The circular and hairpin forms of the Archetypal Loop AL proposed as the "cyclosome". Bottom Middle: A hairpin form of AL. Bottom Right: The most stable hairpin form of the Archetypal Bound AB proposed as a variant of the cyclosome AL.

4. Nucleo-Nucleic and Nucleo-Peptidic Mechanisms

Different intracellular mechanisms involving RNA, DNA, and proteins conserve as relics subsequences of AL, in particular from its short hairpin ATTCAAGATGAAT.

4.1. tRNA Loops

tRNA loops (D-loop, anti-codon loop, T_{Ψ}-loop, and articulation loop) form a sequence that has many similarities to AL. For example, loops of mitochondrial GlytRNA^{GCC} of *Lupine* [46] fit AL almost perfectly (Figure 4) and this tRNA exists in 242 species in the NCBI Nucleotide database [59].

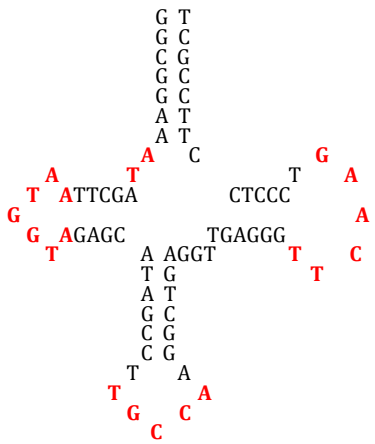

Figure 4. GlytRNA^{GCC} of *Lupine* [66], whose loops (articulation, D-, anti-codon, and T_{Ψ}-loops) fit AL almost perfectly with the sequence formed by its loops (in red).

In the tRNADB-CE database, a high percentage of tRNAs have loops that fit the AL, with TGGTA in D-loop and TTCNA in T_{Ψ}-loop among tRNAs with NTGCCAN as the anticodon loop (Table 2).

Table 2. Percentages of tRNAs containing TGGTA and TTCNA in their D- and T_{Ψ}-loops, among tRNAs having NTGCCAN as the anticodon loop in different species of the tRNADB-CE database.

Species	Percentages
Archaea	248/584 = 42.5%
Bacteria	131983/155823 = 84.7%
Plant	44/80 = 55%
Fungi	106/115 = 92.2%
Virus	6/18 = 33.3%
Phage	67/276 = 24.3%
Chloroplast	109/116 = 94%

4.2. Giant Viruses

The hypothesis that de novo template-free RNAs appear spontaneously—as at the origin of life—and invade modern genomes (in particular those related to the giant viruses) is based on their resemblance to the 25 putative ancestors of the present tRNAs (cf. Figure 3 and [67,68]). Moreover, the AL-pentamers are often observed in the sequences of the giant viruses. To quantify the frequency of the AL-pentamers, we define an AL-proximity frequency for a given genome as the percentage of occurrence in this genome of the 9 most frequent pentamers from the AL (Table 2): If this genome contains 1,000,000 nucleotides, the percentage of such occurrences supposed to be random equals 0.88 ± 0.016* (* for the 90%-confidence interval). From calculations using the NCBI nucleotide database [67–69], the AL-proximity of the complete genome of numerous giant viruses are: Mega 1.82, Mega Chilensis 1.72, Tupan 1.65, Moumou 1.91, Pitho 2.19, Fausto 1.50, Marseille 1.39, Senegal 1.54, Mimi 2.20, Mama 1.90, Bodo 1.81, Samba 1.89, their mean value being equal to 1.80 and that of their virophages Sputnik and Zamilon to 2.23 (see Supplementary Material 1 for other virophages).

4.3. Circular RNAs

The 3801 human circular RNAs from circBase [70] observed after the first discovery of circular RNAs in many organisms [71,72] contain 36228352 possible pentamers; the number of AL-pentamers from a branch of its hairpin form are given in Table 3, which significantly exceeds the number obtained at random.

Table 3. The most frequent pentamers in 3801 human circular RNAs from circBase. The observed numbers can be compared to the number of pentamers obtained at random, 35,379 ± 310* (* for the 90%-confidence interval).

AL-Pentamer	Observed Number
ATTCA	43,219 *
TTCAA	51,917 *
TCAAG	44,233 *
CAAGA	46,523 *
AAGAT	43,189 *
AGATG	48,717 *
GATGA	34,600
ATGAA	51,794 *
TGAAT	44,410 *

4.4. Ribozymes

An RNA catalytic domain has been found within the sequence of the 359 base long negative-strand satellite RNA of tobacco ringspot virus [73]. The catalytic domain contains 2 minimal sequences of satellite RNA, a 14-base substrate RNA, and a 50-base catalytic RNA containing 2 AL-pentamers:

5′-AAACAGAGAAGUCAACCAGAGAAACACACGUUG**UGGUA**UAUUACC**UGGUA**-3′

A minimal RNA hairpin ribozyme discovered 18 years later [74] shows an interesting catalytic activity due to its chain D with 3 AL-pentamers present in its 19 bases: 5′-UCG**UGGUAC**AUUAC**CUGCC**-3′. The AL-tetramer UGGU is generally not cleavable by ribozymes [75], this empirical fact explaining its survival in present ribozymes. AL–pentamers can also be found in the D Chain of many other hairpin ribozymes [76–83], used to build simple RNA systems, consisting of two ribozymes with concerted activity allowing replication [84].

4.5. Intron-Exon Frontier

The heptamers GGTAAGT and TTCA(G)GA present in AL ring are observed frequently at the frontiers of, respectively, exon/intron and intron/exon in genome of many organisms (Figure 5) [85].

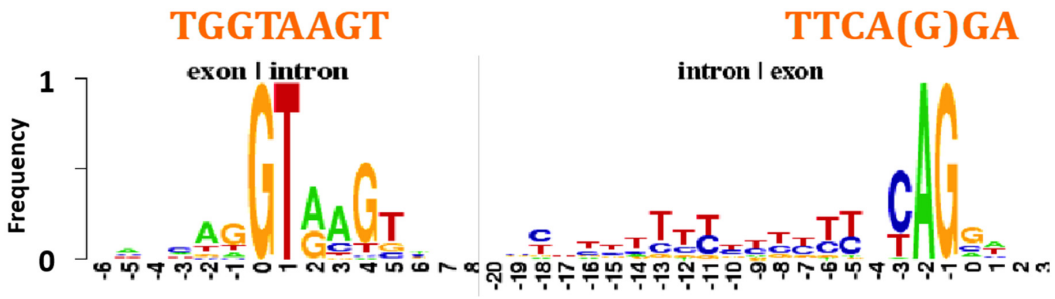

Figure 5. Consensus sequences at exon/intron and intron/exon boundaries in eukaryotes (after [85]).

4.6. Synthetases

Using the AL-proximity calculated from the 9 most frequent pentamers from the AL ring (Table 3), glycyl-tRNA synthetases from [69] have a value more than the 95%-confidence upper threshold, which is equal to 0.88 + 0.49 = 1.37* (calculated for a sequence of size 1000).

Table 4 and Figure 6 show the values of the AL-proximity (for the 9 most frequent AL pentamers) for the tRNA synthetases of the different microorganisms studied in [86–93], especially Bacteria, Archaea, and one Fungus. We observe that Archaea have the maximal values of this proximity and by comparing the sequences of these synthetases [93], we see that the clustering tree based on the sequence resemblance (for the Hamming distance) described as narrow are synthetases having similar values of their AL-proximity (except the pair Haloferax larsenii / Helicobacter pylori).

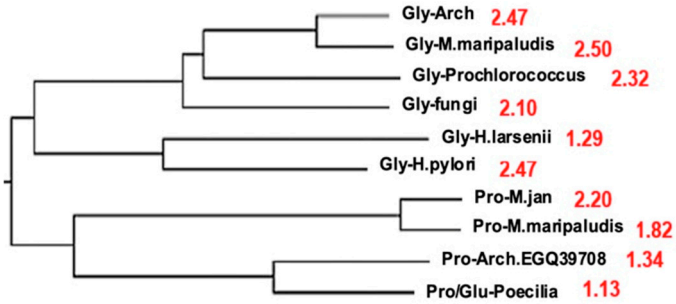

Figure 6. Clustering tree of the synthetases of certain microorganisms based on their sequence resemblance (for the Hamming distance) (after [93]) with indication (in red) of AL-proximity.

Table 4. Gly-tRNA synthetases of different microorganisms [69] with (in red) their AL-proximity.

Homo sapiens mRNA for glycyl-tRNA synthetase, complete cds GenBank: D30658.1 36 × 100/2279 = 1.58 Helicobacter pylori B38 complete genome, strain B38 NCBI Reference Sequence: NC_012973.1: c941829-940933 glycyl—tRNA ligase subunit alpha 15x100/893 = 1.68
Methanococcus maripaludis, strain DSM 2067 chromosome, complete genome NCBI Reference Sequence: NZ_CP026606.1:782166-783890 glycyl-tRNA synthetase 32x100/1721 = 1.86
Fusarium oxysporum f. sp. melonis 26406 unplaced genomic scaffold supercont1.3, whole genome shotgun sequence GenBank: JH659331.1: c2101550-2098309 glycyl-tRNA synthetase 68x100/3238 = 2.10
Prochlorococcus marinus str. NATL1A, complete genome GenBank: CP000553.1: 728949-731111 glycyl-tRNA synthetase 50x100/2155 = 2.32
Methanocaldococcus jannaschii DSM 2661, complete genome GenBank: L77117.1: 219104-220837 glycyl-tRNA synthetase 41x100/1730 = 2.37
Archae Methanococcoides methylutens MM1, complete genome GenBank: CP009518.1: 1554853-1556598 glycyl-tRNA synthetase 43x100/1738 = 2.47
Archae Candidatus Nanosalinarum sp. J07AB56 genomic scaffold scf_7180000039101, whole genome shotgun sequence GenBank: GL982569.1 571934-572506 prolyl-tRNA synthetase 32x100/1286 = 2.49
Methanococcus maripaludis C5, complete genome GenBank: CP000609.1: c1395811-1394084 glycyl-tRNA synthetase 43x100/1720 = 2.50
Methanobacterium formicicum strain BRM9, complete genome GenBank: CP006933.1 556622-558343 glycyl-tRNA synthetase 44x100/1718 = 2.56
Rickettsia prowazekii strain Naples-1chromosome, complete genome GenBank: CP014865.1: c1072366-1071497 glycine-tRNA ligase subunit alpha 25x100/866 = 2.89

The Table 5 gives the values of the AL-proximity for different tRNA synthetases (called also tRNA ligases) and ribosomal or transfer RNAs, showing that the 40S ribosomal RNAs; tRNA synthetases; and 60S, 18S, and 16S ribosomal RNAs have, in this order, decreasing proximities to the AL. The high values of the AL-proximity for the synthetases are consistent with a very early role in a protein-synthesizing machine, by increasing the efficacy of amino acid binding to an RNA-oligopeptide complex such as the AL ring coupled with ligases.

Table 5. The tRNAs of *Thalassiosira pseudonana* (from https://www.ncbi.nlm.nih.gov/nuccore/) with subsequences from AL (left) and AL-proximities (right) in red.

TCAATCAAGATGAAGAGTACGT tRNA synthetase CCMP1335 XM_002286706.1 **2.73**
AGAGTCAAGATGAATAGTAGTA glycyl-tRNA synthetase CCMP1335 XM_002286964.1 **2.44**
CCATGCAAGATGAATGTGGGTG glycyl-tRNA synthetase CCMP1335 XM_002288084.1 **1.92**
GCATTCAAGATGAATCTTCTTG arginyl-tRNA synthetase CCMP1335 XM_002288460.1 2.19
TCCATCTCATGGAATGGTACTG methionyl-tRNA synthetase CCMP1335 XM_002292549.1 **1.91**
CTACCTAGGATGAAGGGTCATG valyl-tRNA synthetase CCMP1335 XM_002295439.1 **2.32**
GCATACAAGAGTAATGGATCTG cysteinyl-tRNA synthetase CCMP1335 XM_002286789.1 **2.04**
CCATTCGAAATGTTTGGTATTG tRNA-Gly mitochondrion DQ186202.1 **7.14**
CCATTGGTGTTGTATGGTAAAC 60S ribosomal protein CCMP1335 XM_002290416.1 **1.83**
CCAAGGAGGATGCGCGAGACTG 60S ribosomal protein CCMP1335 XM_002290087.1 1.77
CTAGTCAAGATGCCTCGTCTAG 40S ribosomal protein CCMP1335 XM_002290013.1 2.78
AAATTGAAGATTAGTGGTGGAG 40S ribosomal protein CCMP1335 XM_002293773.1 **2.97**
CCATGAATGTTTCATGCCTCTG 18S ribosomal protein Bc6EHU KP201658.1 1.55
ACGTTCAACCACACTGGAACTG 16S ribosomal protein BFB575 KC545746.1 **1.51**
CCATTCAAGATGAATGGTACTG CONSENSUS

4.7. Small Acid-Soluble Spore Proteins (SASPs)

SASPs are DNA-binding proteins protecting the DNA backbone from chemical and enzymatic cleavages. Calculating their AL-proximity for all the 22 pentamers of AL (the value in the case of random occurrences of pentamers being equal to 2.1 ± 1.4*) for 100 Bacteria and Archaea randomly chosen in the NCBI database (see Supplementary Material 2) shows that 100% of them have values of AL-proximity over the 95%-confidence upper threshold 2.4*, the mean value being 4.49.

4.8. Defence Mechanisms

The CRISPR-CAS system provides bacteria like *Streptococcus agalactiae* with adaptive immunity and the AL-pentamers ATGGT and ATTCA, and AL-hexamers AATGGT and TCAAGAT (corresponding respectively to the D-loop and Tψ−loop of many ^{t}RNAs) are often found at many levels of the system (CAS proteins, Casposon TIR and CRISP repeats [94]), e.g., typical repeat sequences for CRISPR1 and CRISPR3 [95] contain AL-heptamers shared by AL and tRNA loops:

GTTTTTGTACTCTCAAGATTTAAGTAACTGTACAAC (CRISPR1)

GTTTTAGAGCTGTGTTGTTTCGAATGGTTCCAAAAC (CRISPR3),

as well as the sequences of TIR and CRISPR compared in [96,97], a consensus sequence from central part of the murine RSS VκL8, Jß2.6, and Jß2.2 [98–100], and human RSS spacer common for Vh, V328h2, and V328 [101–103] described in Table 6 and also the protein H354 of Mimivirus kasaii [104] (see Supplementary Material 3 for calculations of AL-proximity of CRISPR-CAS proteins).

Table 6. Sequences of elements involved in defence mechanisms compared to AL.

3′-**ATACATCCC(C)TCTTAAGTTCCCTT**-5′ (TIR)
3′-**TTCCATCCC -TCTTAAGTTCGATT**-5′ (CRISPR)
5′-**ATGGTACTG – CCATTCAAGATGA**-3′ (AL)
5′-**GTGATACAG – CCCTTAACAAAAA**-3′ (murine consensus RSS)
5′-**ATTCAACATGAA**-3′ (human RSS spacer)

The probabilities of a match between the AL and sequences of elements involved in defence mechanisms, at different levels of evolution, are the following:

$p = 2.10^{-9}$ for 19 matches (with an insertion) between TIR and CRISPR using the binomial distribution B(1/4,22), $p = 8.10^{-6}$ for 15 anti-matches between AL and CRISPR plus 1 quasi-anti-match G-T using the distribution B(1/4,21) × B(3/8,1),

$p = 7.10^{-4}$ for 13 matches between AL and consensus RSS using the binomial distribution B(1/4,22), $p = 2.10^{-6}$ for 11 matches between AL and RSS spacer using the binomial distribution B(1/4,12).

In Mimivirus, a mechanism similar to CRISPR has been discovered [104], which involves two exonucleases R 350 and R 354 having respectively 3.02 and 3.71 as the values of the AL-proximity:

- Acanthamoeba castellanii mimivirus DNA, nearly complete genome, strain: Mimivirus kasaii GenBank: AP017644.1 457483-459936 R 350 Lambda-type exonuclease, with AL-proximity **3.02**;
- Acanthamoeba castellanii mimivirus DNA, nearly complete genome, strain: Mimivirus kasaii GenBank: AP017644.1 462878-464527 R 354 Lambda-type exonuclease, with AL-proximity **3.71**.

Moreover, the sequences of the genes of these exonucleases contain numerous heptamers like: 5′-**GATGATGAAGATGATGATGAAGAT**-3′ (MIMIVIRE gene H354).

4.9. Mitochondrial D-loop

In [105], the 2D-structure of the mitochondrial D-loop (7S mtDNA) is given with its central AL-octamer and hexamer TACTGCCAGTCAACATGAAT and in false colour the frequency of its bases (Figure 7). This D-loop is conserved among different species and contains putative mitochondrial micro-RNAs, called mito^{2}miRs in [106].

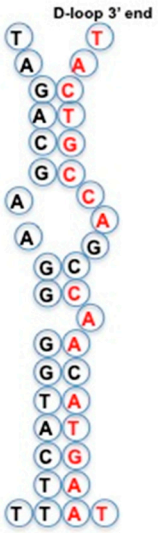

Figure 7. Mitochondrial D-loop (7S mtDNA) with its central AL-octamer and hexamer: TACTGCCAGTCAACATGAAT.

4.10. 5S Ribosomal RNAs

5S ribosomal RNAs are components of the ribosome. Calculating their AL-proximity for all the 22 pentamers of the AL (the value in the case of random occurrences of pentamers being equal to 2.1 ± 1.4*) for 100 Bacteria and Archaea randomly chosen in the NCBI database (see Supplementary Material 4) shows that 78% of them have values of the AL-proximity over the 95%-confidence expected upper threshold of 2.4*, their mean value being equal to 4.01.

4.11. Cytidine Deaminases

The AID/APOBEC protein family comprises cytidine deaminases capable of deaminating cytosine to uracil in the context of a single-stranded polynucleotide [107], met primitively in yeast, and after in fishes, birds, amphibians, and mammals: They play a role of RNA-editing enzymes, contributing to the co-evolution of viruses and their antibodies [108], then perhaps initially to the co-evolution of first RNAs. Among 50 members of this family given in [107] (see Supplementary Material 5), 96% of them

have values of the AL-proximity over the 95%-confidence expected upper threshold of 2.3*, their mean value being equal to 3.36.

5. Discussion

5.1. Origins of the AL Ring

The sequence of the AL ring was obtained by trying to satisfy the constraints of both being as short as possible and being long enough to encode all the amino acids. A justification for the former constraint can be found in the 'lipid world'. Even membranes composed of a single species of molecule can have domains in gel and fluid phases whilst membranes composed of different molecules contain many, predominantly small, domains [109]. We have proposed that such interfaces catalysed the polymerisation of both RNAs and amino acids [110]. In this scenario, in which there would have been many very small domains with closed loop interfaces, there would have been a correspondingly greater production of small RNA rings (Figure 8).

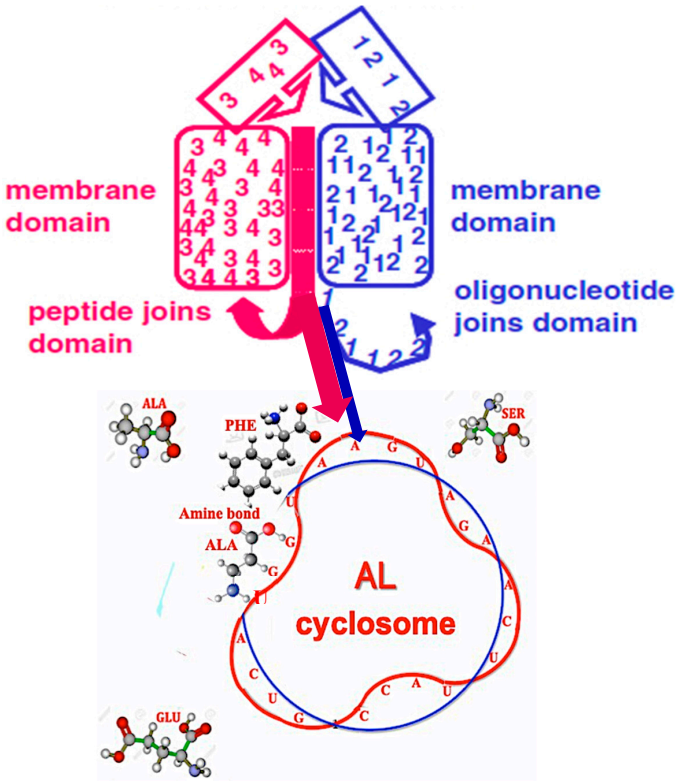

Figure 8. Combination of lipid interfaces mechanism and the functioning of AL as a protein-synthesizing machine without the whole ribosomal machinery.

A justification for the second constraint can be found in the hypothesis that interaction between amino acids and nucleotides stabilised both species thereby leading to their accumulation in the abiotic flux of molecular creation and destruction, as previously proposed [27,28]. In this case, there would have been a strong selection for RNAs to have compositions that would have resulted in the binding of the maximum proportion of the amino acids present in the prebiotic ecology. Hence, if proteins had been synthesised endlessly they would have remained dynamically attached to the selected RNAs and have protected it.

5.2. The AL-Pentamer Proximity as a Marker of Age of the Genome

Class II of aminoacyl-tRNA synthetases constitutes a set of very ancient multi domain proteins [25,93]. By calculating their AL-proximity, we see that their genes are closer to AL than the

genes of the class I (Figure 9 Top). This is available for the 20 synthetases in human, an archaeum (*Methanobacterium lacus*), a proteobacterium (*Rickettsia prowazekii*) close to mitochondria, and an extremophilic bacterium (*Deinococcus radiodurans*, see Supplementary Material 6).

Species	*Homo sapiens*		*Methano-bacterium lacus*		*Rickettsia prowasekii*		*Deinococcus radiodurans*	
class	I	II	I	II	I	II	I	II
a								
AA P	Met 2	Ser 3.6	Met 2	Ser 4.5	Met 3.3	Ser 3.3	Met 1.8	Ser 2.5
	Val 2.1	Thr 3.8	Val 4.5	Thr 3.4	Val 2.7	Thr 3.5	Val 1.2	Thr 2.1
	Leu 3.6	Ala 3.3	Leu 4.8	Ala 2.8	Leu 3.7	Ala 2.6	Leu 1.9	Ala 1.7
	Ile 3.2	Gly 3.2	Ile 3.8	Gly 3.8	Ile 3.9	Gly 3.4	Ile 2	Gly 1.6
	Cys 2.8	Pro 3.1	Cys 4.1	Pro 3.4	Cys 2.5	Pro 3.6	Cys 1.7	Pro 2.5
	Arg 2.6	His 3	Arg 3.6	His 3.4	Arg 2.3	His 2.8	Arg 1.6	His 1.4
b								
	Glu 2	Asp 5.2	Glu 2.8	Asp 5.2	Glu 1.9	Asp 2.7	Glu 2.6	Asp 1.4
	Gln 2.5	Asn 3.3	Gln 2.5	Asn 3.8	Gln 1.6	Asn 2.4	Gln 1	Asn 2.5
		Lys 3.7		Lys 4.1		Lys 4.2		Lys 1.5
c								
	Tyr 2.4	Phe 3	Tyr 3.1	Phe 3.5	Tyr 2.5	Phe 3.2	Tyr 1.5	Phe 2.2
	Trp 2.4		Trp 2.7		Trp 2.4		Trp 1.2	
Mean P	2.56	3.52	3.39	3.79	2.68	3.17	1.67	1.94

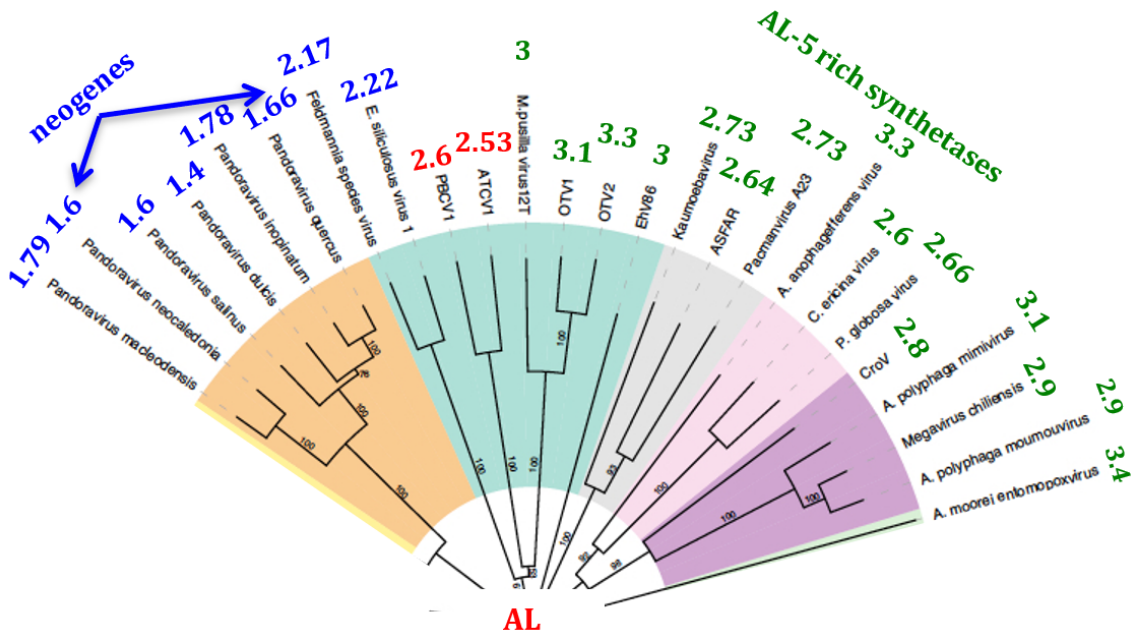

Figure 9. Top: AL-proximity (calculated for the 22 AL-pentamers) of amino-acyl-tRNA synthetases indicated for their classes I and II. Bottom: Giant viruses classification tree. The numbers at the periphery of the circular tree (adapted from [111]) indicate the AL-proximity (calculated for the 9 more frequent AL-pentamers of Table 3) of the Giant viruses' genomes.

The lowest proximity is observed for *Deinococcus radiodurans*, which is capable of genetic transformation by homologous recombination. We observe the same phenomenon for the *Pandora* viruses (Figure 10 Bottom), which are able to create neogenes and which are considered as recently evolved additions to the large family of giant viruses [111]. These observations as well as the order observed between mean AL-proximities of SASPs (4.49), 5S rRNAs (4.01) and cytidine deaminases (3.36), which respectively protect DNA backbone (SASP), act as mediator between tRNA and ribosome

(5S rRNA), and control the cell pyrimidine level (cytidine deaminase) suggest AL-proximity as a marker of genome age, which could constitute a further topic of study.

5.3. 'Infinite' Proteins

The existence of circular mRNAs makes it possible for ribosomes to translate them without ever encountering a translational stop. This could lead to the synthesis of essentially 'infinite' proteins. We propose that the synthesis of such proteins could have occurred at an early stage of the origins of life scenario if the AL cyclosome were simultaneously mRNA/tRNA/synthetase/rRNA. In support of this, the oldest synthetase genes (type II) of *Rickettsia prowazekii* are close to AL (Table 4 and Figure 9 Top), which supports the idea that AL functioned as a primitive protein-synthesizing machine acting without the whole ribosomal machinery for catalysing the first peptides (Figure 10).

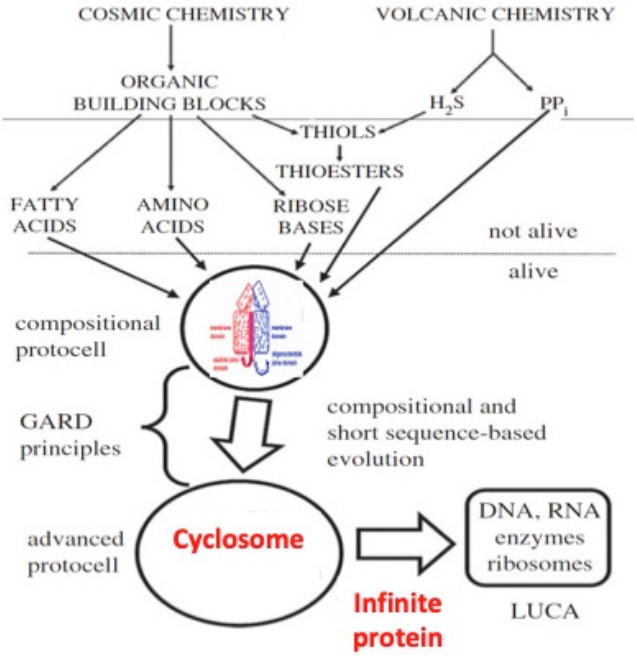

Figure 10. The primitive life machinery (adapted from [92]).

A reversible, stereo-binding between AL and amino acids from a Miller-like source could have catalysed peptide bonds to synthesize a protein with a sequence that would have only been partly random since some juxtaposition, alignment, and orientation on the cyclosome would have occurred [110]; the UGA inside the AL would not necessarily have perturbed the machine because neither reading frames nor punctuation codons would have been needed to produce an "infinite" protein in this way. At a later stage of the evolution of the translational machinery, we propose that synthesis of such proteins would have been associated with (1) a relatively weak primitive Shine-Dalgarno RBS sequence GGAGGU which has a weak complementary sequence inside the AL, CUGCCA, and which would have had the advantage of limiting steric problems due to too many ribosomes trying to bind; (2) a relatively long mRNA; (3) a limited codon repertoire; and (4) the tendency of these proteins to form filaments.

5.4. tRNA Building

A way to build a tRNA molecule from four AL hairpins could consist in following as suggested by many studies [112–114], which propose that the contemporary tRNA was formed by the ligation of four half-sized hairpin-like RNAs. In Figure 11, four partial hairpins from AL ring have been used for

reconstructing the loops of the GlytRNA^{GCC} of *Lupine* [66], this structure having been able to evolve towards the current structure by replacement of unmatched amino acid pairs.

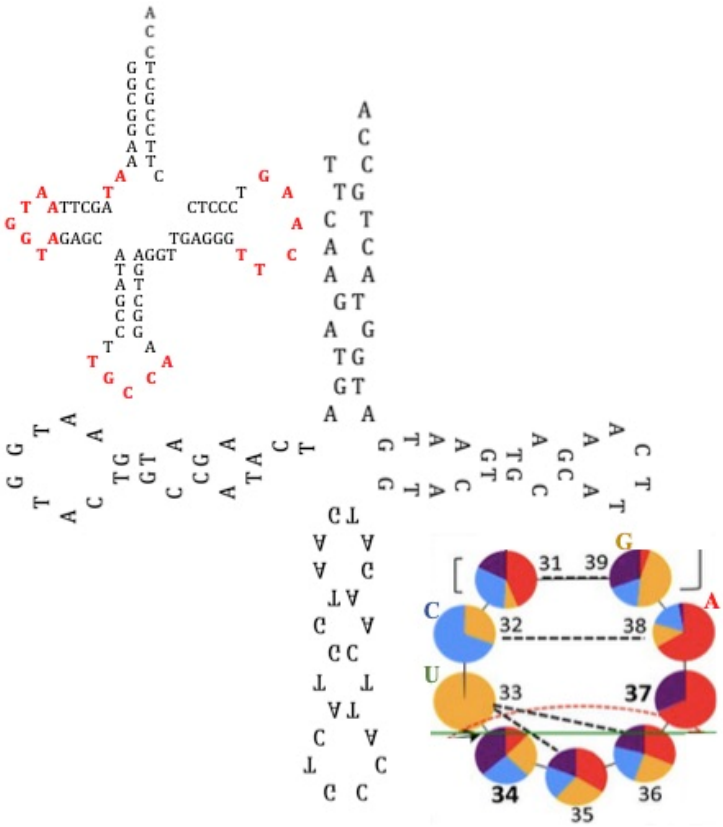

Figure 11. (**a**) The GlytRNA^{GCC} of *Lupine* [66]; (**b**) reconstruction of the loops of the GlytRNA^{GCC} of *Lupine* from four partial hairpins built from AL ring; (**c**) frequency histograms of bases in the anticodon loop of tRNAs belonging to the three domains of life [114].

6. Conclusions

To conclude, a small circular RNA, called AL, has been proposed with a sequence that has the following features:

- Its subsequences (namely, pentamers) are observed as relics in many parts of modern genomes, especially in Archaea;
- AL relics are often present in ^{t}RNA loops, and in mitochondrial D-loops;
- An AL-heptamer constitutes the major part of the exon/intron boundary;
- A scalar proximity to AL explains the relationships between polymerases and, more generally, between complete genomes in phylogenetic trees of Archaea. This proximity suggests a common origin for these genomes.

Hence, the AL cyclosome could have played the role of an ancient protein-synthesizing machine. This claim is central to the stereochemical hypothesis of the genetic code [115] and to the proposal by A. Katchalsky in 1973 [116]: The existence of catalytic RNAs in clays such as the "montmorillonite" may have facilitated the synthesis of small peptides and long RNAs (as is now done by synthetases, polymerases and replicases), thereby constituting an autocatalytic loop at the origin of life.

Hairpin palindromic structure Hairpin size

The existence of a simple RNA structure capable of surviving as a stable hairpin or functioning in a ring form was postulated soon after Katchalsky's hypothesis [46,47,112], and numerous experimental

works [117–119] now reinforce this stereochemical hypothesis in a field that continues to advance both experimentally and theoretically.

We anticipate six research developments will follow from the hypothesis presented here:

- An attempt to take into account the potential evolutional path from the AL ring to the large ribosomal subunits (LSU) extracted from the modular organization of the rRNAs structure [17,94, 120];
- A search for more AL relics in modern genomes at critical functional steps of the nuclear transcription/translation processes (notably when they are coupled as in Archaea [121], in which the Archaea $tRNA^{Gly}$ presents the following sequence in its three successive loops: TGGTA CTGCCA TTCAA, that is a 16-mer from AL [122]), mitochondrial energetic or cellular immune receptor machineries);
- An attempt to explain the evolution of tRNA secondary structures in relation to the genetic code [123–130];
- An attempt to understand the evolution of immune systems (from CRISPR and TOLL to RAG systems [94–97]), taking into account the reuse of former AL RNA fragments already present in the "cyclosome";
- The discovery of sequences linked to AL useful for synthetic biology and studies on "minimal cell" and its primitive genome, with original stable structures as those observed in the "cyclosome" (Figure 12);
- The identification of genetic networks based on common sequences inherited from AL and appearing in regulatory RNAs like microRNAs or circular RNAs.

Hairpin palindromic structure		Hairpin size
C CACTATTCA AGA TGAATGGTG	RNA 3	9
C TACCATTCA AGA TGAATGGTG	RNA 5	9
C TATTCACCA AGA TGGTGAATG	RNA 4	9
C CTATTCACA AGA TGTGAATGG	RNA 6	9
C CTACATTCA AGA TGAATGTGG	RNA 8	9
C TGGTATTCA AGA TGAATGCCA	RNA 7	9
C TGAATGCCA AGA TGGTATTCA	RNA 9	9
C TGCCATTCA AGA TGAATGGTA	AL	9
A TTCATGCCAG A CTGGTATGAA	RNA 1	10
A TTCTGGTATG A CATGCCAGAA	RNA 2	10
A TGTTTATGG AGA CCATAAGCA	RNA 10	9
A TATGTTTGG AGA CCAAGCATA	RNA 11	9
A TATGTTTGG AAG CCAGACATA	RNA 16	9
A CCATAAGCA GAA TGTTTATGG	RNA 14	9
A CCAAGCATA GAA TATGTTTGG	RNA 15	9
A TGTTTATGG AAG CCATAGACA	RNA 18	9
A TGTAGATGG AAG CCATTTACA	RNA 19	9
A TACATTTGG AAG CCAGATGTA	RNA 17	9
A TACAGATGG AAG CCATTTGTA	RNA 20	9
A TATGGTTCT GCA AGAACCATG	RNA 23	9
A TACCATTCT GCA AGAATGGTG	RNA 24	9
A TGAATGGT GCCATTCA AG A CT	AB	10 (1 unpaired base)
T TCATGCCAG AAA CTGGTATGA	RNA 12	9
T TCTGGTATG AAA CATGCCAGA	RNA 21	9
G ATGCCAGAA ACA TTCTGGTAT	RNA 22	9

Figure 12. The palindromic structure of the 25 rings. The couples or triplets of rings are obtained by exchanging the blue and the red subsequences.

Supplementary Materials: The following are available online at http://www.mdpi.com/2075-1729/9/2/51/s1.

Author Contributions: Both authors contributed equally to the research described in the manuscript.

Acknowledgments: We are greatly indebted to J. Besson, A. Henrion-Caude, A. Moreira, R. Thomas, and P. Tracqui for many discussions about the molecular structures potentially involved at the origin of life.

Conflicts of Interest: The authors declare that there is no conflict of interest regarding the publication of this paper.

References

1. Palade, G.E.; Siekevitz, P. Liver microsomes. An integrated morphological and biochemical study. *J. Biophys. Biochem. Cytol.* **1956**, *2*, 171–214. [CrossRef] [PubMed]
2. Wolfram, S. *A New Kind of Science*; Wolfram Media, Inc.: Champaign, IL, USA, 2002.
3. Gardner, M. Mathematical Games: The Fantastic Combinations of John Conway's New Solitaire Game 'Life'. *Sci. Am.* **1970**, *22*, 120–123. [CrossRef]
4. Besson, J.; Gavaudan, P.; Schützenberger, M.P. Sur l'existence d'une certaine corrélation entre le poids moléculaire des acides aminés et le nombre de triplets intervenant dans leurs codages. *C. R. Acad. Sci.* **1969**, *268*, 1342–1344.
5. Demongeot, J.; Hazgui, H. The Poitiers school of mathematical and theoretical biology: Besson-Gavaudan-Schützenberger's conjectures on genetic code and RNA structures. *Acta Biotheor.* **2016**, *64*, 403–426. [CrossRef] [PubMed]
6. Bloch, D.; McArthur, B.; Widdowson, R.; Spector, D.; Guimaraes, R.C.; Smith, J. tRNA-rRNA sequence homologies: A model for the origin of a common ancestral molecule, and prospects for its reconstruction. *Orig. Life* **1984**, *14*, 571–578. [CrossRef]
7. Altman, S. Enzymatic cleavage of RNA by RNA. *Biosci. Rep.* **1990**, *10*, 317–337. [CrossRef] [PubMed]
8. Cech, T.R. Self-splicing and enzymatic activity of an intervening sequence RNA from Tetrahymena. *Biosci. Rep.* **1990**, *10*, 239–261. [CrossRef]
9. Agmon, I.; Auerbach, T.; Baram, D.; Bartels, H.; Bashan, A.; Berisio, R.; Fucini, P.; Hansen, H.; Harms, J.; Kessler, M.; et al. On peptide bond formation, translocation, nascent protein progression and the regulatory properties of ribosomes. *Eur. J. Biochem.* **2003**, *270*, 2543–2556. [CrossRef]
10. Tamura, K.; Schimmel, P. Peptide synthesis with a template-like RNA guide and aminoacyl phosphate adaptors. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 8666–8669. [CrossRef]
11. O'Donoghue, P.; Luthey-Schulten, Z. On the Evolution of Structure in Aminoacyl-tRNA Synthetases. *Microbiol. Mol. Biol. Rev.* **2003**, *67*, 550–573. [CrossRef]
12. Agmon, I.; Bashan, A.; Zarivach, R.; Yonath, A. Symmetry at the active site of the ribosome: Structural and functional implications. *Biol. Chem.* **2005**, *386*, 833–844. [CrossRef] [PubMed]
13. Tamura, K.; Schimmel, P.R. Chiral-selective aminoacylation of an RNA minihelix: Mechanistic features and chiral suppression. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 13750–13752. [CrossRef] [PubMed]
14. Demongeot, J.; Glade, N.; Moreira, A.; Vial, L. RNA relics and origin of life. *Int. J. Mol. Sci.* **2009**, *10*, 3420–3441. [CrossRef] [PubMed]
15. Demongeot, J.; Ben Amor, H.; Gillois, P.; Noual, M.; Sené, S. Robustness of regulatory networks. A Generic Approach with Applications at Different Levels: Physiologic, Metabolic and Genetic. *Int. J. Mol. Sci.* **2009**, *10*, 4437–4473. [CrossRef] [PubMed]
16. Agmon, I. The Dimeric Proto-Ribosome: Structural Details and Possible Implications on the Origin of Life. *Int. J. Mol. Sci.* **2009**, *10*, 2921–2934. [CrossRef]
17. Davidovich, C.; Belousoff, M.; Bashan, A.; Yonath, A. The evolving ribosome: From non-coded peptide bond formation to sophisticated translation machinery. *Res. Microbiol.* **2009**, *160*, 487–492. [CrossRef] [PubMed]
18. Bokov, K.; Steinberg, S.V. A hierarchical model for evolution of 23S ribosomal RNA. *Nature* **2009**, *457*, 977–980. [CrossRef]
19. Szostak, J.W. Origins of life: Systems chemistry on early Earth. *Nature* **2009**, *459*, 171. [CrossRef]
20. Bashan, A.; Agmon, I.; Raz Zarivach, R.; Schluenzen, F.; Harms, J.; Berisio, R.; Bartels, H.; Franceschi, F.; Auerbach, T.; Hansen, H.A.S.; et al. Structural Basis of the Ribosomal Machinery for Peptide Bond Formation, Translocation, and Nascent Chain Progression. *Mol. Cell* **2013**, *11*, 91–102. [CrossRef]
21. Huang, L.; Krupkin, M.; Bashan, A.; Yonath, A.; Massa, L. Protoribosome by quantum kernel energy method. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 14900–14905. [CrossRef]
22. Bernhardt, H.S.; Tate, W.P. A Ribosome Without RNA. *Front. Ecol. Evol.* **2015**, *3*, 129. [CrossRef]

23. Agmon, I. Could a Emerge Spontaneously in the Prebiotic World? *Molecules* **2016**, *21*, 1701. [CrossRef] [PubMed]
24. Krupkin, M.; Wekselman, I.; Matzov, D.; Eyal, Z.; Diskin Posner, Y.; Rozenberg, H.; Zimmerman, E.; Bashan, A.; Yonath, A. Avilamycin and evernimicin induce structural changes in rProteins uL16 and CTC that enhance the inhibition of A-site tRNA binding. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 6796–6805. [CrossRef] [PubMed]
25. Kim, Y.; Kowiatek, B.; Opron, K.; Burton, Z.F. Type-II tRNAs and Evolution of Translation Systems and the Genetic Code. *Int. J. Mol. Sci.* **2018**, *19*, 3275. [CrossRef] [PubMed]
26. Norris, V. Why do bacteria divide? *Front Microbiol.* **2015**, *6*, 322. [CrossRef] [PubMed]
27. Maturana, H.R.; Varela, F.J. *Autopoiesis and Cognition: The Realization of the Living*; Reidel: Boston, MA, USA, 1980.
28. Bourgine, P.; Stewart, J. Autopoiesis and cognition. *Artif. Life* **2004**, *10*, 327–345. [CrossRef] [PubMed]
29. Hunding, A.; Kepes, F.; Lancet, D.; Minsky, A.; Norris, V.; Raine, D.; Sriram, K.; Root-Bernstein, R. Compositional complementarity and prebiotic ecology in the origin of life. *BioEssays* **2006**, *28*, 399–412. [CrossRef] [PubMed]
30. Robinson, R. Jump-starting a cellular world: Investigating the origin of life, from soup to networks. *PLoS Biol.* **2005**, *3*, e396. [CrossRef] [PubMed]
31. Ono, N.; Ikegami, T. Self-maintenance and self-reproduction in an abstract cell model. *J. Theor. Biol.* **2000**, *206*, 243–253. [CrossRef]
32. Ono, N.; Ikegami, T. Artificial chemistry: Computational studies on the emergence of selfreproducing units. In Proceedings of the 6th European Conference on Artificial Life (ECAL'01), Prague, Czech Republic, 10–14 September 2001; Kelemen, J., Sosik, S., Eds.; Springer: Berlin, Germany, 2001; pp. 186–195.
33. Weil, G.; Heus, K.; Faraut, T.; Demongeot, J. An archetypal basic code for the primitive genome. *Theor. Comput. Sci.* **2004**, *322*, 313–334. [CrossRef]
34. Demongeot, J.; Moreira, A. A circular RNA at the origin of life. *J. Theor. Biol.* **2007**, *249*, 314–324. [CrossRef] [PubMed]
35. 5S RNAdb. Available online: http://www.combio.pl/rrna/alignment/ (accessed on 25 March 2019).
36. Kinefold. Available online: http://kinefold.curie.fr (accessed on 25 March 2019).
37. Fonville, N.C.; Velmurugan, K.R.; Tae, H.; Vaksman, Z.; McIver, L.J.; Garner, H.R. Genomic leftovers: Identifying novel microsatellites, over-represented motifs and functional elements in the human genome. *Sci. Rep.* **2016**, *6*, 27722. [CrossRef] [PubMed]
38. Tóth, G.; Gáspári, Z.; Jurka, J. Microsatellites in different eukaryotic genomes: Survey and analysis. *Genome Res.* **2000**, *10*, 967–981. [CrossRef] [PubMed]
39. Pemberton, T.J.; Sandefur, C.I.; Jakobsson, M.; Rosenberg, N.A. Sequence determinants of human microsatellite variability. *BMC Genom.* **2009**, *10*, 612. [CrossRef] [PubMed]
40. Willard, H.F.; Waye, J.S. Chromosome-specific Subsets of Human Alpha Satellite DNA: Analysis of Sequence Divergence Within and Between Chromosomal Subsets and Evidence for an Ancestral Pentameric Repeat. *J. Mol. Evol.* **1987**, *25*, 207–214. [CrossRef]
41. Zhuma, T.; Tyrrell, R.; Sekkali, B.; Skavdis, G.; Saveliev, A.; Tolaini, M.; Roderick, K.; Norton, T.; Smerdon, S.; Sedgwick, S.; et al. Human HMG box transcription factor HBP1: A role in hCD2 LCR function. *EMBO J.* **1999**, *18*, 6396–6406. [CrossRef]
42. Presnyak, V.; Alhusaini, N.; Chen, Y.H.; Martin, S.; Morris, N.; Kline, N.; Olson, S.; Weinberg, D.; Baker, K.E.; Graveley, B.R.; et al. Codon optimality is a major determinant of mRNA stability. *Cell* **2015**, *160*, 1111–1124. [CrossRef]
43. Hobish, M.K.; Wickramasinghe, N.S.; Ponnamperuma, C. Direct interaction between amino acids and nucleotides as a possible physicochemical basis for the origin of the genetic code. *Adv. Space Res.* **1995**, *15*, 365–382. [CrossRef]
44. Yarus, M.; Widmann, J.J.; Knight, R. RNA-amino acid binding: A stereo chemical era for the genetic code. *J. Mol. Evol.* **2009**, *69*, 406–429. [CrossRef]
45. Yarus, M. The Genetic Code and RNA-Amino Acid Affinities. *Life* **2017**, *7*, 13. [CrossRef]
46. Polyansky, A.A.; Zagrovic, B. Evidence of direct complementary interactions between messenger RNAs and their cognate proteins. *Nucleic Acids Res.* **2013**, *41*, 8434–8443. [CrossRef] [PubMed]
47. Zagrovic, B.; Bartonek, L.; Polyansky, A.A. RNA-protein interactions in an unstructured context. *FEBS Lett.* **2018**, *592*, 2901–2916. [CrossRef] [PubMed]

48. Demongeot, J. Sur la possibilité de considérer le code génétique comme un code à enchaînement. *Rev. De Biomaths* **1978**, *62*, 61–66.
49. Demongeot, J.; Besson, J. Code génétique et codes à enchaînement I. *C. R. Acad. Sci. Série III* **1983**, *296*, 807–810.
50. Demongeot, J.; Besson, J. Genetic code and cyclic codes II. *C. R. Acad. Sc. Série III* **1996**, *319*, 520–528.
51. Demongeot, J.; Drouet, E.; Moreira, A.; Rechoum, Y.; Sené, S. Micro-RNAs: Viral genome and robustness of the genes expression in host. *Philos. Trans. R. Soc. A* **2009**, *367*, 4941–4965. [CrossRef] [PubMed]
52. Demongeot, J.; Glade, N.; Moreira, A. Evolution and RNA relics. A Systems Biology view. *Acta Biotheoretica* **2008**, *56*, 5–25. [CrossRef] [PubMed]
53. Demongeot, J.; Hazgui, H.; Bandiera, S.; Cohen, O.; Henrion-Caude, A. MitomiRs, ChloromiRs and general modelling of the microRNA inhibition. *Acta Biotheor.* **2013**, *61*, 367–383. [CrossRef] [PubMed]
54. Demongeot, J.; Cohen, O.; Henrion-Caude, A. MicroRNAs and Robustness in Biological Regulatory Networks. A Generic Approach with Applications at Different Levels: Physiologic, Metabolic, and Genetic. In *Systems Biology of Metabolic and Signaling Networks*; Aon, M.A., Saks, V., Schlattner, U., Eds.; Springer: Berlin, Germany, 2013; pp. 63–114.
55. GtRNAdb. Available online: http://lowelab.ucsc.edu/GtRNAdb/Rhod_spha_ATCC_17029/rhodSpha_ATCC17029-tRNAs.fa (accessed on 25 March 2019).
56. Müller, M.; Legrand, C.; Tuorto, F.; Kelly, V.P.; Atlasi, Y.; Lyko, F.; Ehrenhofer-Murray, A.E. Queuine links translational control in eukaryotes to micronutrient from bacteria. *Nucl. Acids Res.* **2019**, *47*, 3711–3727. [CrossRef] [PubMed]
57. Choi, H.; Otten, S.; McClain, W.H. Isolation of novel $tRNA^{Ala}$ mutants by library selection in a $tRNA^{Ala}$ knockout strain. *Biochimie* **2002**, *84*, 705–711. [CrossRef]
58. Agris, P.F.; Vendeix, F.A.P.; Graham, W.D. tRNA's Wobble Decoding of the Genome: 40 Years of Modification. *J. Mol. Biol.* **2007**, *366*, 1–13. [CrossRef] [PubMed]
59. Weixlbaumer, A.; Murphy, F.V.; Dziergowska, A.; Malkiewicz, A.; Vendeix, F.A.P.; Agris, P.F.; Ramakrishnan, V. Mechanism for expanding the decoding capacity of transfer RNAs by modification of uridines. *Nat. Struct. Mol. Biol.* **2007**, *14*, 498–502. [CrossRef] [PubMed]
60. Targanski, I.; Cherkasova, V. Analysis of genomic tRNA sets from Bacteria, Archaea, and Eukarya points to anticodon–codon hydrogen bonds as a major determinant of tRNA compositional variations. *RNA* **2008**, *14*, 1095–1109. [CrossRef] [PubMed]
61. Fergus, C.; Barnes, D.; Alqasem, M.A.; Kelly, V.P. The Queuine Micronutrient: Charting a Course from Microbe to Man. *Nutrients* **2015**, *7*, 2897–2929. [CrossRef] [PubMed]
62. Choi, H.; Gabriel, K.; Schneider, J.; Otten, S.; McClain, W.H. Recognition of acceptor-stem structure of $tRNA^{Asp}$ by Escherichia coli aspartyl-tRNA synthetase. *RNA* **2003**, *9*, 386–393. [CrossRef] [PubMed]
63. Root-Bernstein, R. Simultaneous origin of homochirality, the genetic code and its directionality. *BioEssays* **2007**, *29*, 1–10. [CrossRef] [PubMed]
64. Root-Bernstein, R.; Root-Bernstein, M. The ribosome as a missing link in prebiotic evolution II: Ribosomes encode ribosomal proteins that bind to common regions of their own mRNAs and rRNAs. *J. Biol.* **2016**, *397*, 115–127. [CrossRef]
65. Root-Bernstein, R.; Root-Bernstein, M. The Ribosome as a Missing Link in Prebiotic Evolution III: Over-Representation of tRNA- and rRNA-Like Sequences and Plieofunctionality of Ribosome-Related Molecules Argues for the Evolution of Primitive Genomes from Ribosomal RNA Modules. *Int. J. Mol. Sci.* **2019**, *20*, 140. [CrossRef]
66. Bartnik, E.; Borsuk, P. A glycine tRNA gene from lupine mitochondria. *Nucleic Acids Res.* **1986**, *14*, 2407. [CrossRef]
67. Seligmann, H.; Raoult, D. Stem-Loop RNA Hairpins in Giant Viruses: Invading rRNA-Like Repeats and a Template Free RNA. *Front Microbiol.* **2018**, *9*, 101. [CrossRef]
68. Seligmann, H.; Raoult, D. Unifying view of stem–loop hairpin RNA as origin of current and ancient parasitic and non-parasitic RNAs, including in giant viruses. *Curr. Opin. Microbiol.* **2016**, *31*, 1–8. [CrossRef] [PubMed]
69. NCBI. Available online: https://blast.ncbi.nlm.nih.gov/ (accessed on 25 March 2019).
70. Circbase. Available online: http://www.circbase.org/ (accessed on 25 March 2019).

71. Ivanov, A.; Memczak, S.; Wyler, E.; Torti, F.; Porath, H.T.; Orejuela, M.R.; Piechotta, M.; Levanon, E.Y.; Landthaler, M.; Dieterich, C.; et al. Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Rep.* **2014**, *10*, 170–177. [CrossRef] [PubMed]
72. Liang, T.; Yang, C.; Li, P.; Liu, C.; Guo, L. Genetic analysis of loop sequences in the let-7 gene family reveal a relationship between loop evolution and multiple isomiRs. *PLoS ONE* **2014**, *9*, e113042. [CrossRef] [PubMed]
73. Hampel, A.; Tritz, R. RNA Catalytic Properties of the Minimum (-)sTRSV Sequence. *Biochemistry* **1989**, *28*, 4929–4933. [CrossRef] [PubMed]
74. Salter, J.; Krucinska, J.; Alam, S.; Grum-Tokars, V.; Wedekind, J.E. Water in the Active Site of an All-RNA Hairpin Ribozyme and Effects of Gua8 Base Variants on the Geometry of Phosphoryl Transfer. *Biochemistry* **2006**, *45*, 686–700. [CrossRef]
75. Pérez-Ruiz, M.; Barroso-delJesus, A.; Berzal-Herranz, A. Specificity of the Hairpin Ribozyme. *J. Biol. Chem.* **1999**, *274*, 29376–29380. [CrossRef] [PubMed]
76. Müller, U.F. Design and Experimental Evolution of trans-Splicing Group I Intron Ribozymes. *Molecules* **2017**, *22*, 75. [CrossRef]
77. Paul, N.; Joyce, G.F. A self-replicating ligase ribozyme. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12733–12740. [CrossRef]
78. Perreault, J.; Weinberg, Z.; Roth, A.; Popescu, O.; Chartrand, P.; Ferbeyre, G.; Breaker, R.R. Identification of Hammerhead Ribozymes in All Domains of Life Reveals Novel Structural Variations. *PLoS Comput. Biol.* **2011**, *7*, e1002031. [CrossRef]
79. Hammann, C.; Luptak, A.; Perreault, J.; De La Peña, M. The ubiquitous hammerhead ribozyme. *RNA* **2012**, *18*, 871–885. [CrossRef]
80. Harris, K.A.; Lünse, C.E.; Li, S.; Brewer, K.I.; Breaker, R.R. Biochemical analysis of hatchet self-cleaving ribozymes. *RNA* **2015**, *21*, 1–7. [CrossRef] [PubMed]
81. Rupert, P.B.; Ferré-D'Amaré, A.R. Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature* **2001**, *410*, 780–786. [CrossRef] [PubMed]
82. Gebetsberger, J.; Micura, R. Unwinding the twister ribozyme: From structure to mechanism. *WIREs RNA* **2017**, *8*, e1402. [CrossRef] [PubMed]
83. Chapple, K.E.; Bartel, D.P.; Unrau, P.J. Combinatorial minimization and secondary structure determination of a nucleotide synthase ribozyme. *RNA* **2003**, *9*, 1208–1220. [CrossRef] [PubMed]
84. Luisi, P.L. *The Emergence of Life: From Chemical Origins to Synthetic Biology*; Cambridge University Press: Cambridge, UK, 2006.
85. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*; Garland Science: New York, NY, USA, 2002.
86. Hartman, H. Speculations on the evolution of the genetic code I. *Orig. Life* **1975**, *6*, 423–427. [CrossRef]
87. Hartman, H. Speculations on the evolution of the genetic code II. *Orig. Life* **1978**, *9*, 133–136. [CrossRef] [PubMed]
88. Hartman, H. Speculations on the evolution of the genetic code III: The evolution of t-RNA. *Orig. Life* **1984**, *14*, 643–648. [CrossRef]
89. Hartman, H. Speculations on the origin of the genetic code IV. *J. Mol. Evol.* **1995**, *40*, 541–544. [CrossRef]
90. Hartman, H.; Favaretto, P.; Smith, T.F. The archaeal origins of the eukaryotic translational system. *Archaea* **2006**, *2*, 1–9. [CrossRef]
91. Smith, T.F.; Lee, J.C.; Gutell, R.R.; Hartman, H. The origin and evolution of the ribosome. *Biol. Direct* **2008**, *3*, 16. [CrossRef]
92. Hartman, H.; Smith, T.F. The evolution of the ribosome and the genetic code. *Life* **2014**, *4*, 227–249. [CrossRef] [PubMed]
93. Smith, T.F.; Hartman, H. The evolution of Class II Aminoacyl-tRNA synthetases and the first code. *FEBS Lett.* **2015**, *589*, 3499–3507. [CrossRef] [PubMed]
94. Lier, C.; Baticle, E.; Horvath, P.; Haguenoer, E.; Valentin, A.S.; Glaser, P.; Mereghetti, L.; Lanotte, P. Analysis of the type II-A CRISPR-Cas system of Streptococcus agalactiae reveals distinctive features according to genetic lineages. *Front. Genet.* **2015**, *6*, 214. [CrossRef] [PubMed]
95. Horvath, P.; Romero, D.A.; Coûté-Monvoisin, A.C.; Richards, M.; Deveau, H.; Moineau, S.; Boyaval, P.; Fremaux, C.; Barrangou, R. Diversity, Activity, and Evolution of CRISPR Loci in Streptococcus thermophilus. *J. Bacteriol.* **2008**, *190*, 1401–1412. [CrossRef] [PubMed]

96. Koonin, E.V.; Krupovic, M. Evolution of adaptive immunity from transposable elements combined with innate immune systems. *Nat. Rev. Genet.* **2015**, *16*, 184–192. [CrossRef] [PubMed]
97. Koonin, E.V.; Makarova, K.S. Mobile Genetic Elements and Evolution of CRISPR-Cas Systems: All the Way There and Back. *Genome Biol. Evol.* **2017**, *9*, 2812–2825. [CrossRef] [PubMed]
98. Lee, A.I.; Fugmann, S.D.; Cowell, L.G.; Ptaszek, L.M.; Kelsoe, G.; Schatz, D.G. A Functional Analysis of the Spacer of V(D)J Recombination Signal Sequences. *PLos Biol.* **2003**, *1*, 56–69. [CrossRef] [PubMed]
99. Pasqual, N.; Gallagher, M.; Aude-Garcia, C.; Loiodice, M.; Thuderoz, F.; Demongeot, J.; Ceredig, R.; Marche, P.N.; Jouvin-Marche, E. Quantitative and Qualitative Changes in ADV-AJ Rearrangements During Mouse Thymocytes Differentiation: Implication for a Limited TCR ALPHA Chain Repertoire. *J. Exp. Med.* **2002**, *196*, 1163–1174. [CrossRef] [PubMed]
100. Thuderoz, F.; Simonet, M.A.; Hansen, O.; Dariz, A.; Baum, T.P.; Hierle, V.; Demongeot, J.; Marche, P.N.; Jouvin-Marche, E. Numerical Modelling of the V-J Combinations of the T Cell Receptor TRA/TRD Locus. *PLoS Comp. Biol.* **2010**, *6*, e1000682. [CrossRef] [PubMed]
101. Ramsden, D.A.; Baetz, K.; Wu, G.E. Conservation of sequence in recombination signal sequence spacers. *Nucl. Acids Res.* **1994**, *22*, 1785–1796. [CrossRef]
102. Takeuchi, N.; Ishiguro, N.; Shinagawa, M. Molecular cloning and sequence analysis of bovine T-cell receptor gamma and delta chain genes. *Immunogenetics* **1992**, *35*, 89–96. [CrossRef] [PubMed]
103. Liu, M.F.; Robbins, D.L.; Crowley, J.J.; Sinha, S.; Kozin, F.; Kipps, T.J.; Carson, D.A.; Chen, P.J. Characterization of four homologous L chain variable region genes that are related to 6B6.6 idiotype positive human rheumatoid factor L chains. *J. Immunol.* **1989**, *142*, 688–694.
104. Levasseur, A.; Bekliz, M.; Chabrière, E.; Pontarotti, P.; La Scola, B.; Didier Raoult, D. MIMIVIRE is a defence system in mimivirus that confers resistance to virophage. *Nature* **2016**, *531*, 249–252. [CrossRef] [PubMed]
105. Portugene. Available online: http://fpereira.portugene.com/research1.html (accessed on 25 March 2019).
106. Woese, C. The biological significance of the genetic code. *Prog. Mol. Subcell. Biol.* **1969**, *1*, 5–46.
107. Conticello, S.C.; Thomas, C.J.F.; Petersen-Mahrt, S.K.; Neuberger, M.S. Evolution of the AID/APOBEC Family of Polynucleotide (Deoxy)cytidine Deaminases. *Mol. Biol. Evol.* **2005**, *22*, 367–377. [CrossRef]
108. Martinez, T.; Shapiro, M.; Bhaduri-McIntosh, S.; MacCarthy, T. Evolutionary effects of the AID/APOBEC family of mutagenic enzymes on human gamma-herpesviruses. *Virus Evol.* **2019**, *5*, vey040. [CrossRef]
109. Kraft, M.L.; Weber, P.K.; Longo, M.L.; Hutcheon, I.D.; Boxwer, S.G. Phase separation of lipid membranes analyzed with high-resolution secondary ion mass spectrometry. *Science* **2006**, *313*, 1948–1951. [CrossRef]
110. Raine, D.J.; Norris, V. Lipid domain boundaries as prebiotic catalysts of peptide bond formation. *J. Theor. Biol.* **2007**, *246*, 176–185. [CrossRef]
111. Legendre, M.; Fabre, E.; Poirot, O.; Jeudy, S.; Lartigue, A.; Alempic, J.M.; Beucher, L.; Philippe, N.; Bertaux, L.; Christo-Foroux, E.; et al. Diversity and evolution of the emerging Pandoraviridae family. *Nat. Commun.* **2018**, *9*, 2285. [CrossRef]
112. Di Giulio, M. On the Origin of Protein Synthesis: A Speculative Model Based on Hairpin tRNA Structures. *J. Theor. Biol.* **1994**, *171*, 303–308. [CrossRef]
113. Tamura, K. Origins and Early Evolution of the tRNA Molecule. *Life* **2015**, *5*, 1687–1699. [CrossRef] [PubMed]
114. Grosjean, H.; Westhof, E. An integrated, structure- and energy-based view of the genetic code. *Nucleic Acids Res.* **2016**, *44*, 8020–8040. [CrossRef] [PubMed]
115. Fontecilla-Camps, J.C. The Stereochemical Basis of the Genetic Code and the (Mostly) Autotrophic Origin of Life. *Life* **2014**, *4*, 1013–1025. [CrossRef] [PubMed]
116. Paecht-Horowitz, M.; Katchalsky, A. Synthesis of amino acyl-adenylates under prebiotic conditions. *J. Mol. Evol.* **1973**, *2*, 91–98. [CrossRef] [PubMed]
117. Meyer, S.C.; Nelson, P.A. Can the origin of the genetic code be explained by direct RNA templating? *Bio-Complex.* **2011**, *2011*, 1–10. [CrossRef]
118. Lancet, D.; Zidovetzki, R.; Markovitch, O. Systems protobiology: Origin of life in lipid catalytic networks. *J. R. Soc. Interface* **2018**, *15*, 20180159. [CrossRef]
119. Hsiao, C.; Mohan, S.; Kalahar, B.K.; Williams, L.D. Peeling the onion: Ribosomes are ancient molecular fossils. *Mol. Biol. Evol.* **2009**, *26*, 2415–2425. [CrossRef]
120. Opron, K.; Burton, Z.F. Ribosome Structure, Function, and Early Evolution. *Int. J. Mol. Sci.* **2018**, *19*, 40. [CrossRef]

121. French, S.L.; Santangelo, T.J.; Beyer, A.L.; Reeve, J.N. Transcription and translation are coupled in Archaea. *Mol. Biol. Evol.* **2007**, *24*, 893–895. [CrossRef]
122. Pak, D.; Du, N.; Kim, Y.; Sun, Y.; Burton, Z.F. Rooted tRNAomes and evolution of the genetic code. *Transcription* **2018**, *9*, 137–151. [CrossRef]
123. Seligmann, H. Giant viruses as protein-coated amoeban mitochondria? *Virus Res.* **2018**, *253*, 77–86. [CrossRef] [PubMed]
124. Seligmann, H. Protein sequences recapitulate genetic code evolution. *Comput. Struct. Biotechnol. J.* **2018**, *16*, 177–189. [CrossRef]
125. Seligmann, H. Alignment-based and alignment-free methods converge with experimental data on amino acids coded by stop codons at split between nuclear and mitochondrial genetic codes. *Biosystems* **2018**, *167*, 33–46. [CrossRef] [PubMed]
126. Demongeot, J.; Seligmann, H. Theoretical minimal RNA rings recapitulate the order of the genetic code's codon-amino acid assignments. *J. Theor. Biol.* **2019**, *471*, 108–116. [CrossRef]
127. Demongeot, J.; Seligmann, H. Spontaneous evolution of circular codes in theoretical minimal RNA rings. *Gene* **2019**, *705*, 95–102. [CrossRef] [PubMed]
128. Demongeot, J.; Seligmann, H. Bias for 3′-dominant codon directional asymmetry in theoretical minimal RNA rings. *J. Comput. Biol.* **2019**, 26. [CrossRef] [PubMed]
129. Demongeot, J.; Seligmann, H. More pieces of ancient than recent theoretical minimal proto-tRNA-like RNA rings in genes coding for tRNA synthetases. *J. Mol. Evol.* **2019**, *87*, 1–23. [CrossRef] [PubMed]
130. Torres de Farias, S.; Gaudêncio Rêgo, T.; José, M.V. Origin of the 16S Ribosomal Molecule from Ancestor tRNAs. *Sci* **2019**, *1*, 8. [CrossRef]

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).