THE DYNAMICS OF THE LIVING

Romeu Cardoso Guimarães

Laboratório de Biodiversidade e Evolução Molecular, Departamento de Genética,
Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, 31270-901 Belo Horizonte MG Brazil.
TelFax (+55-31)-32744988, Mobile -988976439
romeucardosoguimaraes@gmail.com , romeucg@icb.ufmg.br

In: **Systemics, Self-Organization and Information**, Eds. Alfredo Pereira Jr, Chris Nunn, William Pickering, Ricardo Gudwin. Submitted to Routledge UK, November 2017

Abstract Living beings are metabolic flow systems that self-construct on the basis of memories, and adapt/evolve on the basis of constitutive plasticity. Life is the ontogenetic and evolutionary process instantiated by living beings. Biological flow dynamics is continuous with universal dynamics. Cellular flow initiates at the search-and-capture behaviors of receptors for nutrients. It is centered on the system of synthesis of amino acids and their trapping in proteins. The latter generated the nucleic acid template system that works as sequence memories. Protein diversification constructs the flow structures and catalysts. Continuous activity of the protein synthesis sink is guaranteed by the reduction of protoplasmic mass at the shedding of vesicles and at the fission that inaugurates the reproductive and evolutionary cycles. Evolution is inherent to vital dynamics. It is driven by challenges originated from external variability and from a series of stressful conditions that are inherent to organic constitution or created by organisms themselves – including environmental degradation – that induce repair and regenerative activities.

Key-words Genetic code; Origins of life; Self-reference; tRNA dimers; Universal flow; Metabolic flow; Metabolic sink; Reproductive pull; Protein plasticity; Replicative memory.

Graphical Abstract is Figure 2 (Graphical abstracts have become usual in contemporary journals; I don't know if the book will adopt them)

List of topics	
Introduction	Spontaneity and plasticity
Overview	The drive and its endogenization
The code and metabolism	Conflicts, tension, contradictions
The sink plus plasticity	Specificity starts with the genetic code
The pre-RNA world	Memory in strings
Coherence-decoherence	Cycles of repetitiveness
Life and living beings	Membranes
Intriguing convergences	Exploratory behaviors
Living and the general evolution	Construction of the cellular flow system
Self-reference, self-stimulation, and auto-catalysis	Lincoln, Maxwell
Geochemical scenario	Accumulated mistakes and lesions
Dimers of oligomeric proto-tRNAs	Good and bad
The intracellular sink	Reproduction
Activities in the proto-cellular RNP globule	Behaviors focus on the tips of the flow system

Introduction It is not easy to reach consensus about concepts of nature, the categorization of living beings, and the life process. Some researchers have even suggested that these goals are not worth the dedication of practicing experimentalist scientists, that scientists can proceed in their work without those particular concepts, and that the search for clear definitions of them is a waste of time and effort [5]. Statements like

these may be taken as quasi-jokes, however. It is only to be expected that fundamental questions will always be present in the minds of scientists, and as soon as any interesting possibility of an answer arises such questions will receive attention and interest on the part of researchers.

Our own search is in accord with the perspective that we should not deceive our audience, who are no doubt expecting a contribution to the subject. 'What is life?' is a question of general interest, and one that is of particular interest to students and textbook authors. The study of evolution, general and biological, needs a firm grounding in the molecular sciences, centered on the transition from geochemical systems to biological ones. A characterization of early cells is useful in defining what is original and what is derived, and what is accreted in subsequent evolutionary processes.

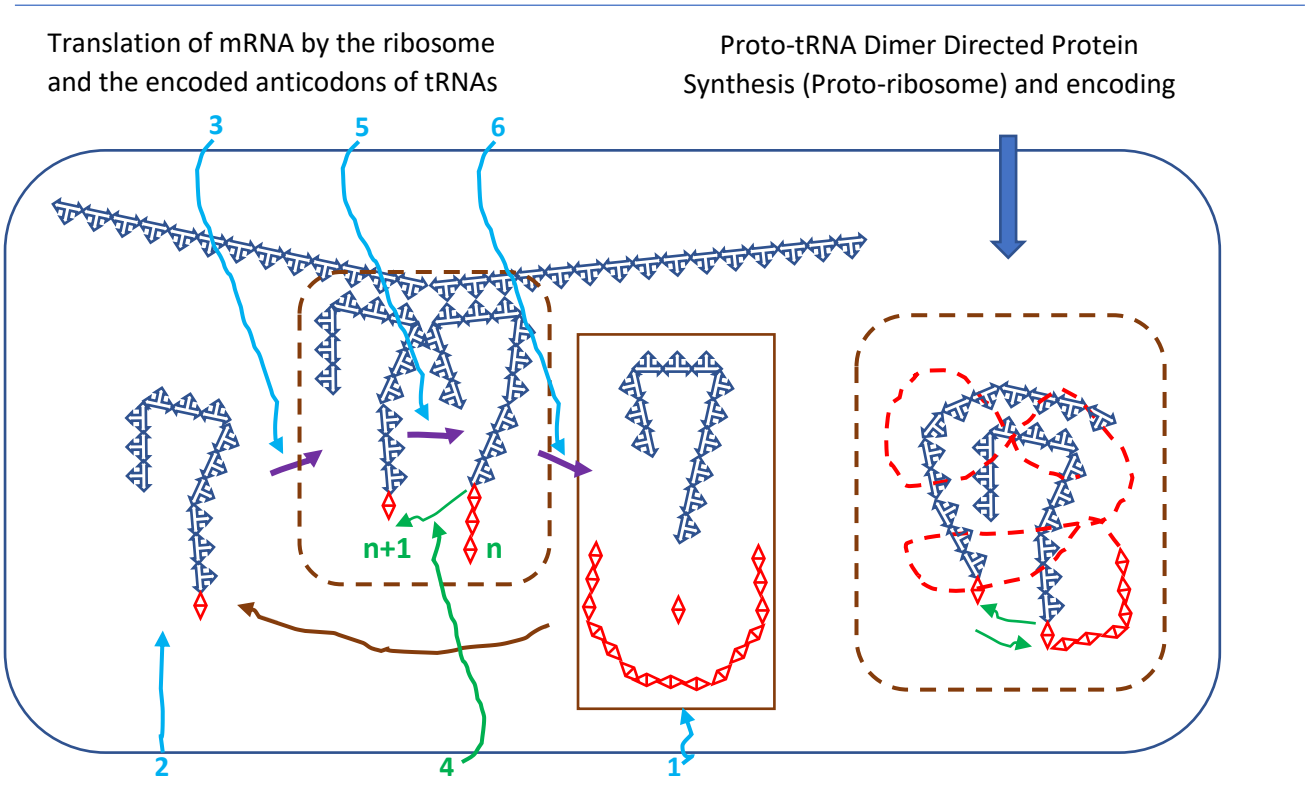

Figure 1. **Decoding the chain of codons by the encoded tRNAs in the ribosomal machine of cells (left)**. (1) The protein enzyme aminoacyl-tRNA synthetase unites the cognate pair of amino acid (red diamond) and tRNA, forming (2) an aminoacyl-tRNA that (3) enters the ribosome. Through the transferase reaction (4), the aminoacyl-tRNA receives the peptidyl, with n amino acids, from the peptidyl-tRNA, and the n+1 new peptidyl-tRNA is (5) translocated to its site, so that the tRNA that donated the peptidyl n exits from the ribosome and (6) may be charged in a new cycle. **The proto-ribosome (right)**. In Dimer-Directed Protein Synthesis the anticodons of the dimerized (proto)tRNAs are at the same time codons for each other, and the transferase reaction is adirectional or bidirectional, there being no position fixed for the amino acid or the nascent peptide. The peptide product binds back to and coats the producer tRNA(s). When this precursor-product system reaches stability and productivity, encoded pairs (cognate tRNA and synthetase) are formed.

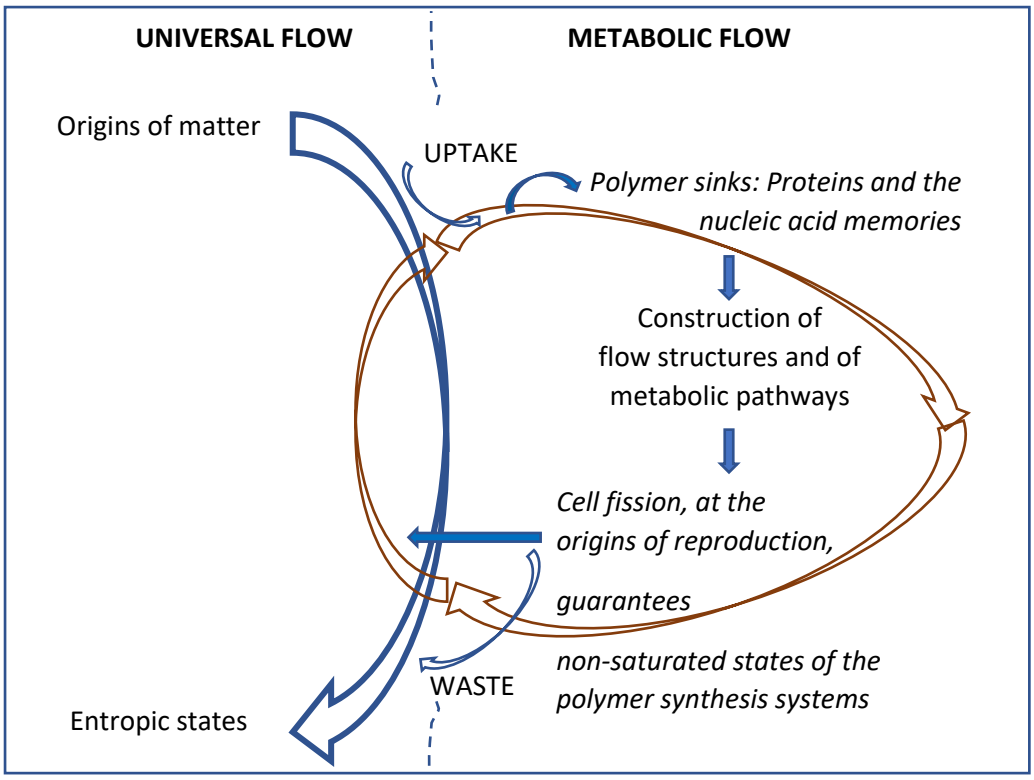

Figure 2. **The universal flow system includes the biological metabolic flow system**. The environment is degraded through loss of material taken up by the organisms that expel waste into it. The internal drive of living beings is the sink/trap constituted by the synthesis of polymers. These construct a variety of flow-directing structures. The sink is maintained continuously active through extrusion of protoplasm chunks, thereby avoiding saturation. The most sensitive flow structures are at the extremities: uptake by the nutritional receptors and the extrusion of chunks. The latter is most salient in the cell fission that takes place at reproduction. The psychic analogs are, respectively, the desires for food and for reproduction or sex.

Overview The Self-Referential Model (**SRM**) for the origins of the genetic code suggests a succession of events in the formation of the cell [9]. It starts with protein synthesis directed by dimers of oligomeric proto-tRNAs. The protein products bind to their producers in a self-referential loop, and stabilize the ensemble that is a proto-ribosomal RNP, a protein production system (Figure 1). Metabolic pathways start with utilization of C1 substrates to synthesize Glycine and Serine, the first amino acids encoded. Encoding is the first instance of biological specificity. The (Gly)Ser Cycle of Methylotrophs contains precursors to gluconeogenesis from which the central metabolic pathways develop, including nucleic acid biosynthesis. The vital dynamics thus configured is that of a metabolic flow system (Figure 2) which is **(1)** centered on the protein synthesis sink, and which has to be constantly kept active and healthy. Nucleic acids, aside from their original role as producer machinery, develop the ligation of codes into replicative memories (genes) for protein sequences. Protein diversification then builds **(2)** cellular structures that are directed to guarantee the flow, which starts with uptake of substrates from the environment and includes the extrusion of waste into the environment, which is degraded. This is only one among a variety of challenges and stressful conditions that organisms have to face from external and internal sources which take part in forcing the evolution of the flow system. The cells can only answer with further diversification. Their resources for this reside in the plasticity of their components – more extensive in proteins than in RNA, the least in DNA – and of the network organization. The model for the protocell is that of a 'spongy RNP globule' in water, which

maintains the internal/external distinction through protein spontaneous motility and binding activities. **(3)** The globule cannot grow beyond a certain limit where aggregation and surface tension forces are overcome and chunks are lost. This introduces a stimulatory effect on the flow system, guaranteeing that the protein synthesis activity is kept active. Such spontaneous benefit is combined with the solution to the problem of extrusion of waste. This is problematic with respect to nitrogenous compounds, which cannot be transformed into to gases like the carbon waste. The problem is greater with proteins whose catabolism may leave behind some indigestible remnants that are toxic and form entangled aggregates. Their elimination as chunks are a necessity. These associated benefits are at the origin of the ubiquitous cellular character of shedding vesicles and exosomes. **(4)** The energetic cost associated with such losses of protoplasm were partially circumvented when some of them received portions of the genetic memories and became daughter cells at the beginning of the reproductive cycling. Accomplishment of this regulated protoplasmic fission marks both the acquisition of potentially perennial activity of the protein synthesis sink and the installation of the Darwinian process [7].

The code and metabolism The study of the genetic code yielded information on the successive steps of amino acid incorporation into the biosystem. A repetitive modular structure was configured by networks of pairs of anticodon triplets. Modularity facilitated and speeded up the encoding process. It pointed to Glycine and Serine as the first encodings, the first module. This replaces the hypothesis of the origins of the first amino acids encoded being the two most abundant prebiotics, Glycine and Alanine. Otherwise, they indicated the start from the anabolic pathway of gradual stepwise incorporation of one-carbon units (C1) directly into the C2-amino acid Glycine and C3 Serine. The Glycine-Serine pathway of some Methylotrophic prokaryotes [1] contains this subset – possibly its original core – and adds other C3 compounds of the Pyruvate family, the largest being in the C4 range, Oxaloacetate and Malyl-CoA.

Such configuration makes this pathway the simplest in the central metabolism; it is the only pathway containing amino acids, not just precursors to them, and it also comprises precursors to the gluconeogenesis pathways. This set of activities comprises a starting point for an early cell. The novelty of this approach is the central focus on amino acids and their incorporation into protein synthesis, which is a proto-ribosomal function. After that appear the early protein functions beyond the ribosomal, which should include the synthesis of amino acids (Gly and Ser, and others thereafter) and the aminoacyl-(proto)tRNA synthetase.

C1 is at the bottom of the well with respect to substrate size, and has a flexibility of choices all along the series from the most reduced CH4 to the most oxidized CO2. There is no way to go deeper or to obtain more guaranteed non-fluctuating sources. Assimilation processes obtain C2 compounds from the exogenous C1, as acetate derivatives, beyond the Gly and the C3 Ser pathway. The Gly ⇔ Ser interconversion is accomplished by one of the most flexible enzymes, Serine HydroxyMethyl Transferase [4]. C2 compounds also feed bioenergetics with Acetyl-CoA, and directly feed the biosynthesis of lipids.

The development of gluconeogenesis, and thereafter of the glycolytic route, the pentose-phosphate shunt, and the citrate cycle, completed the majority of the central metabolism, from which nucleic acids could then be built. Amino acids are necessary components of the pathways of biosynthesis of the nucleobases – Aspartate for the pyrimidines and Glycine for purines – to the point that both the proteins and nucleobases can be considered the end-points (sinks) of the amino acid biosynthesis pathways. The catabolic pathway of the biosynthesis of Gly and Ser, which became evolutionarily more prevalent, follows from C3 phosphorylated sugars, of the glycolytic pathway, to phosphoserine, and then finally to Ser and Gly.

Sink plus plasticity A molecular system proposal corresponding to the long-searched-for '*vital force*' is now equated to a series of metabolic sink-effective mechanisms. The old term expressed the necessity of understanding 'what would be inside the living system' that could be responsible for its apparent impetus

and innate tendency to demonstrate spontaneity and initiative with respect to itself and to its environments, both physicochemical and sociobiological.

For such characteristics that enable living beings to develop agency, we utilize two concepts. One is the simple *sink mechanism* centered on amino acids, with their sequestration in proteins and their utilization for biosynthesis of other biomolecules and for bioenergetics. The Thioester world proposal for bioenergetics – of Christian De Duve and more recently Bill Martin [22] – is most consistent with this 'C1 World' scenario. A summarizing statement says that the biological core is the protein synthesis system, which includes the genetic system as memory, as regulator of memory expression and as template coder for the synthesis of proteins. This core should be kept continually active, or at least fully able to resume activity at any moment.

For the second concept we adopt the term *plasticity*, with a somewhat broadened meaning. Plasticity is among the main properties of proteins, and is what mediates the resulting adaptations and evolution of organisms. The strategy for this choice is parsimonious: maximize what is already available and minimize the introduction of terminological and conceptual novelties, which would require more effort for explanation and understanding. Some phenotypic adaptations that reach stability (in the sense of becoming repetitive) in the expression network may destabilize some nodes in that network through overuse, to the point of provoking genetic variability within them and thus becoming mutational hotspots upon which selection may work. One well-known destabilization mechanism is the coupling between transcription, replication, and repair. Highly transcriptionally active genes may be damaged to the point of activating replication and repair mechanisms, and some of these are mutation prone.

The pre-RNA world Our approach does not require a start from a rich RNA world. In its early stages, protein synthesis would have been directed by pairs of prebiotic oligomers that accomplished the functions of present-day tRNAs inside ribosomes. While the ribosomal pairs are associated laterally, the tRNAs being positioned according to the directional mRNA chain of codons, the prebiotic proto-ribosomes would be composed of only the two proto-tRNAs associated by base pairing of the proto-anticodon sites (Figure 1). This arrangement would be inside the heterogeneous category of Non-Ribosomal Protein Synthesis (NRPS) systems [6,15,16], called Dimer-Directed Protein Synthesis (DDPS). There is some similarity of this with the known case of NRPS directed by an aminoacyl-tRNA synthetase fragment that sequentially takes two aminoacyl-tRNAs to synthesize a peptide.

The constitution of the proto-tRNA is not known, there being ample room for choosing among diverse possibilities. The oligomers might be synthesized from minerals, whose repetitiveness is analogous to replication, there being no need for sophisticated replication mechanisms as developed by the RNA techniques. The role of replication in the template realm is of utmost importance for the identity and stability of the whole system, but its memory role indicates the concomitance of its origin with that of proteins, to whom the memories are directed. The SRM is, therefore, Proto-RNP or RNP from the beginning.

In the same way as the constitution of the producer oligomers is left open, so are the substrates that they carry and utilize for becoming part of the products. The dimer-directed polymerization mechanism would possibly work generically with different kinds of substrates. With so many unknowns, we have to delineate experiments that would illuminate the road from geochemistry to the RNP world.

Coherence-decoherence The DDPS process has some peculiarities that are worth being analyzed in themselves and compared with the ribosome- and mRNA-directed processes. The (proto)tRNA associations are dynamic, via hydrogen bonding, and may generate different states: (a) complementary with other (proto)tRNAs, forming the dimers, which opens the route to DDPS; (b) complementary with other RNAs, which may open the route toward translation of mRNA; (c) binding to proteins, which inaugurates the RNP associations such as the ribosomal and the aminoacyl-tRNA synthetases at encoding.

State (a) is called coherent or superposed, where the components are of the same kind and may present undecided identities and functions. The transferase activity is adirectional or bidirectional, and the codon and anticodon functions are interchangeable. This would have been the only one present at the initial encoding times, during the formation of the LUCA populations. The two proto-tRNAs are mutually equivalent: their proto-anticodon sites accomplish the codon function for one another at the same time. This situation is analogous to the quantum mechanics superposition of states in a quantum object. In this case, a wavepacket would have components in a coherent or undecided state, and may probabilistically produce the classical wave or the particle states (and associated properties) after passing through the interactions that lead to decoherence, including those that are part of the detection or measurement processes. The dimer is also adirectional or bidirectional in the sense that any of the partners may serve the aminoacyl- or peptidyl-carrier functions, and may exchange the functions in each round of the realization of the transferase function, which is a job of the joint pair.

States (b, c) are decohered and each (proto)tRNA may acquire individuality as 'classical' components of the cellular translation machinery. The transition from the DDPS to the ribosome- and mRNA-directed state would involve the intromission of a decohering interaction. This could occur, among various possibilities, with (c) the peptide products of the DDPS that may be heterogeneous and able to bind differentially to the oligomers, or with the entry of another oligomer in the place of one of the members of the dimer, which would be taking the role of the classical mRNA.

Manfred Eigen's group proposed tRNA to be 'the first gene' to be translated, which is a variant of (b) and requires a previous encoding mechanism, such as (a). A consensus RNY triplet constitution was obtained for the tRNA sequence collection in the 1970's. This suggestion for the constitution of early codes has been explored by Edward Trifonov [21], Marco José [11] and Sávio Farias [3]. A subset of the RNY-coded amino acids is the GNC Val Ala Gly Asp, abundant among prebiotic syntheses, studied by Kenji Ikehara [10]. In case tRNAs would be involved with both encoding via dimerization and with the mRNA function, extreme self-reference would be obtained.

Life and living beings A short conceptual statement will set the stage for the discussion on the vital dynamics:

> ***Living beings*** *are metabolic flow systems that*
> *self-construct on the basis of memories and*
> *adapt / evolve on the basis of constitutive plasticity.*
> ***Life*** *is the ontogenetic and evolutionary process instantiated by living beings*

Explication of some key-words: **A**. Living beings are easier to define due to being concrete objects. **B**. Life is more difficult to define, due to its being a process with many abstract components. Living is the activity of both creating and exhibiting the life process. **C**. Metabolism refers to creation of new molecular combinations through chemical transformations, and it is relational. It implies dependency on the environment that is the source of substrates and provides the ultimate sink for disposal of waste. **D**. The flow indicates that a healthy system should have the transformations running smoothly and unimpeded; if this doesn't happen, the system gets sick. **E**. Memories for self-construction indicate the properties of repetition and regeneration of components. The mechanisms of memories are not detailed here for reasons of space. There are the memories in strings, the genes, and memories in self-sustained cycles that are typical of epigenetics. **F**. Plasticity is the property of structures that are capable of changing states in accordance with environmental or internal conditions. Some of these states may participate in different epigenetic mechanisms, thereby generating possible fixation of the correspondent genetic characters and this leading to evolutionary properties. Plasticity is inherent and extensive in protein 3D dynamics, less intense in RNA, and even less in DNA.

The idea of the flow is not new, but we add a plausible rational explanation for it, which spans from the entropic universe to the origins of life and to reproduction. In medical genetics, the idea of flow is essential to the concept of the Inborn Errors of Metabolism. In the Darwinian account, reproductive flow is measured as fitness. We now generalize the flow concept (its wider application to the classification of diseases is being developed) and pinpoint its centrality to protein synthesis.

Intriguing convergences The most salient aspects of dynamics, with stronger appeal to the general observer, are **(I)** the relational and interactive behaviors, at the openings of the metabolic mechanisms to the environments, mostly at the uptake domains, that is, feeding, and **(II)** the reproductive drive. It is tempting to highlight the counterparts of these in the most evident among human psychological characteristics: the desires, impulses, and drives for food and sex. We show further down how the metabolic flow concept accommodates these two aspects in a consistent rationale.

The convergence is noteworthy between the description of human characters by psychologists and those attributes that can be discerned the basis of biochemical reasoning. Instead of just coincidental final results of investigations, the mutual fit may mean that our minds, as indicated by diverse realms of study, would be following a tendency or a bias in favor of repetitions of mechanisms. It is possible that the background to these constancies is engrained in the natural selection mechanisms that would be continually forcing the adjustments and adaptations between organisms and environments; our minds would also be biased in this direction. In the ethological and psychological realms, they would be reflected in some 'cognitive architectures' of all minds, configured like the Jungian archetypes.

Evidence for these coincidences arise repetitively during our studies, intriguingly enough to raise suspicions of some kind of constraining or directedness in our thoughts and reasonings. I was first alerted to this when finding that the initial protein conformations indicated by the SRM were the Intrinsically Disordered segments [9,20]. This is consistent with the quantum mechanics rationale that their objects – wavepackets – are also disordered. In both cases, the order – reflected in informational patterns – would arise at the interaction of entities that were originally disordered [8]. The same mechanism shows up in very different realms of study. It seems that our minds can only be relaxed, pleased, and happy, when some kind of 'informational closure' is reached. The alternative would be instabilities and loose ends in our lines of reasoning, something that would lead to intellectual tension.

The informational closure concept was applied when a relatedness was found between the formation of the initiation and termination mechanisms at translation of mRNA. The SRM indicates that the entire set of elongation codes was formed utilizing a 'primitive punctuation system' based only on the higher metabolic stability of the protein head segments and the lower stability in the tail segments. The last codes introduced were the specific punctuation, starting with addition of one specific anticodon for initiation. The system immediately deleted the anticodons that were in conflict with that initiation triplet, whose codon complements became the terminators.

In favor of the convergence being real, not just a bias, there is the highly prevalent (and justifiable within the scientific community) principle of parsimony: 'Multiplicity in explanations is acceptable only when there is good evidence'.

The living and the general evolution Our study centers on the polymer (nucleoprotein) constitution of living beings. In our view, the prebiotic world reached only the oligomer level of polymerization. It departs from the contemporary mainstream trend by attributing logical primacy to proteins, not to RNA. The mechanism of origin places both kinds of polymers together (RNP World instead of RNA World) in mutuality and in coevolution.

Cells are not autonomous entities, because they reach only partial sustainability. The SRM requires a pre-existing dynamical process referring to the prevailing universal evolution model. The paradigm of evolutionary self-organization of matter is followed. It goes along the downhill gradients of energies that instantiate interactions between atoms and molecules. New forms are created, and in the interactions inside them a fraction of the energies is dampened but another fraction remains exposed, modified as new interactive radicals and as free energies from which self-organization and evolution are maintained.

Self-reference is only one of the processes inside the general process of self-organization. It considers that even in a simple two-body interaction that forms a new composite body, there exists at the least the mutuality where, at this limit, the bodies depend on (refer to) one another interchangeably, so that the notions of self and non-self dissolve and fuse together, and the one/other distinction makes no sense. This kind of theoretical conundrum may not apply to the case of a system observed by a biochemist, where the two components (nucleic acids and proteins) are clearly distinguishable. The same applies to molecular transformations that are partly external (the substrates taken up and the waste extruded by the organisms) or internal to cells/organisms (the metabolic products). Other compounds remain, belonging to and staying the same in both compartments, as, for example, the mineral atoms and the components of water.

Self-reference, self-stimulation, and auto-catalysis In the SRM, self-reference means the formation of the producer-product bidirectional loop, where the two directions accomplish different functions, at the creation of a self-stimulated or self-feeding production system. The mechanism is akin or analogous to auto-catalytic systems or sets.

This idea may be made easily accessible to the layman's understanding by thinking of it as analogous to any artisanship or industrial setup that is attempting to become a successful production system, exemplifying learning as a natural evolutionary process. In an enterprise, the producer develops products and offers them to the population of eventual buyers. The producer will only continue producing those products that find buyers (the consumers in the market system), whose positive feedback informs the producer with positive stimuli. The valuable initiative (choice, information) is in the hands of the products themselves. The producers may only submit variety for the consumer's decisions. In case these do not feedback positively, the producer stops making them and tries other alternatives.

The proto-tRNA dimer, through its transferase activity, produces peptides. These products, through their intrinsic stability and backward-binding activity, create a proto-RNP composite structure. This may continue manifesting producer activity, which is now the work of a *bona fide* peptide producing system. The prolonged cycling of the system is a result of the stability demonstrated by the peptides, which should be evolutionarily modulated – strong enough to promote the cycling and duration of activities, but not so strong as to risk hampering the activity. This mode of promotion of stability is equivalent to systemic self-stimulation or self-feeding.

Geochemical scenario Experiments on the abiotic synthesis of organic compounds utilizing simple conditions such as electric discharges (among other energy sources), boiling water, and gases, that simulate what would have constituted Earth's early oceans and atmosphere (different concentrations of ammonia, carbon monoxide or dioxide, hydrogen, alone or with added inorganic ions or metals) have been conducted by many groups in the last few decades. Other natural 'experiments' have been followed, such as those 'conducted' in or around crustal hot springs, volcanoes, and submarine hot vents, and on the samples collected from meteorites, comets, and interstellar dust.

The results vary mostly quantitatively, but agree largely in the abundance of mainly organic acids and a variety of amino acids, including some of those naturally occurring in proteins. Abundance under those conditions means also stability or reduced chemical reactivity. Cyclic compounds are also among the less

reactive, due to the absence of exposed radicals at the tips of molecular chains. In meteorites, nucleobases have been found in some thousand-fold less concentrations, parts-per-billion, in contrast to the amino acids that reach parts-per-million. Given the extensive variety of possibilities, one expects lots of cycles of evolutionary mutual adjustments among components that became the biological choices for proteins and nucleic acids.

Human feelings are not sympathetic to the entropic direction, but it can be evaluated less unpalatably by looking at it from the following perspectives: (a) It is just a physical model, not a truth (this is an idealist concept, not a scientific one); (b) The entropic states are infinitely far in the future, and are not realistically within the dimensions of human mesoscopic ranges; (c) In human dimensions, there is plenty of room for changes in the model and for possible future developments, given that that the cultural/evolutionary trajectory is essentially open.

The overall mechanism: (a) involves a downhill trajectory of gradients of mass and energy; (b) includes the 'reduction of turbulence' in the system with the association of elements in higher order organizations (molecules) with mutual dampening of free energies; (c) forms, shapes and activities are exposed as patterns that observers utilize as information. Chemical evolution scenarios present many turbulence-provoking compounds with excesses of exposed free radicals or energies that would get involved with many reactions and instabilities. One of the modes for reaching stability was the amination of the organic acids. Among these, the alpha-keto acids gave origin to the natural amino acids, their detoxified and more stable derivatives.

It happens that such serial detoxification reactions do not show signs of reaching a bottom line. The ongoing evolution testifies to the evidence that most reactions may only partially dampen the reactivities. Aside from the partial reductions, other reactivities may arise with unforeseen results. These observations take part in the concepts of evolution being open-ended and non-directional. There are only a few examples of relatively neutral and non-reactive compounds that could extend the category of the atomic noble gases or of some other 'dark' matter or energy.

Dimers of oligomeric proto-tRNAs The constitution of the oligomers that compose the dimers is not known. There are many alternatives simpler than the RNA, such as pre- or proto-RNAs with different sugars or substitutes for the phosphate. A favorite route for the repetitive synthesis of oligomers, as a substitute for replication, is the activity of crystals (surfaces and borders) such as those present in clays. A condition that would make easier the transition to RNA is that of the oligomers already containing the nucleobases or similar compounds. An interesting example is the PNA – Peptide Nucleic Acid – that has a poly-glycine-like backbone that holds the nucleobases at the correct intervals for a nucleic acid mimic but, due to this backbone not being acidic, makes very tight double-helices, the strands being difficult to separate. Due to this rigidity, it is not considered a good candidate for a prebiotic oligomer. Electrostatic repulsion of the phosphate (acidic) backbones is of help to the flexibility of the double-helices in the water environment.

Experimental examination of the transferase activity of dimers cannot, obviously, be conducted on the (unknown) prebiotic oligomers, but may find a proxy or substitute in the present-day RNAs. It is suggested that tests could start inside the variety of known mini-tRNAs or mini-helices. These should receive segments with the ability of base-pairing. The ability of these mini-tRNAs to be spontaneously aminoacylated is well known, so that the dimer experiments are expected to occur in the near future.

The intracellular sink Construction of the cellular bodies depends mostly on the carbon and nitrogen compounds; most abundant are the amino acids of proteins, and after that the nucleobases of nucleic acids. Secondarily enter oxygen and hydrogen, partly in sugars and lipids, partly associated with nitrogenous

compounds, plus the sulfur in some amino acids and the phosphate in nucleic acids. The sum leads to the CHONPS acronym.

The polymers are considered traps or sinks of the monomers, which are derived from simple elements taken up from the environment. The SRM says that the main source compounds for the origins of cells are amino acids, from which sugars and nucleobases are derived. Some of these more complex molecules were available from geochemical processes but would have participated mostly in some pre- or proto-biotic mechanisms. Their quantities would have been submitted to fluctuations that impeded the construction of cells on their bases. In the case of amino acids, only Glycine coincides in both realms: it is abundant abiotically and the first in biosynthesis.

Activities in the proto-cellular RNP globule When the peptides produced by the proto-tRNA dimers bind to the producers, the resulting RNP forms a globule in water. The continued activity of the RNP system is supported by the uptake of amino acids from the solution where the globule is embedded, and by oligomers derived from crystal-mediated synthesis. The globule is structured as a spongy gel. Its constituent polymers are submitted to continuous water-ionic-thermal stress. The constitutive polymers, especially the more flexible peptides, exhibit continuous trembling or vibrational dynamics, which is provoked by their interaction with the polar radicals of water. Some of the interactions may even result in hydrolysis, which is among the main intrinsic tensions/contradictions inside biosystems: water is necessary, but at the same time a challenge and a source of stress to be counteracted; it is expelled during polymerization, but these develop in a watery environment. This 'dialectical' challenge has, otherwise, the 'good' facet of provoking variations that contribute to evolutionary dynamics.

Such movements inside the sponge characterize the in/out distinction. The inside may be considered 'warmer' than the environment, in consequence of the polymer agitation. The binding activities inside the globule create gradients that provoke *facilitated diffusion*. Materials taken in will be rarefied in the pools at the periphery of the globule, which will be replenished by simple diffusion. These movements along gradients are maintained as long as the trapping activity in the globule is present.

Spontaneity and plasticity These movements, most typical of proteins-in-water, are the beginnings of what is called the essential spontaneity of living systems. A more usual term for this character is plasticity. It is more frequently applied to phenotypes, the pleomorphisms of the bodily appearances, but is now extended to the epigenetic properties.

The drive and its endogenization The stage is set for selection, among variants of the constituents of the globule, for those that contribute to permanence and growth. The process may reach situations close to the self-feeding capacity, through substitution of the out-sourcing by endogenous mechanisms. The process is called endogenization. The drive for permanence and growth is initially the general evolution of the external flow: materials taken in contribute to *reduction of the turbulence* outside, so that the stability, overall and of the globule, is favored.

The drive is implemented by endogenization, where some metabolic pathways being created make the trap and sink mechanism more effective. A portion of the set of components of the gradient mechanism becomes synthesized inside the globule. Some external mechanisms are partially substituted by internal ones. The mechanisms are made more efficient and less sensitive to external fluctuations at the beginning of the creation of what is called, in cells or organisms, the *internal milieu*.

Conflicts, tension, contradictions The general evolutionary flow is drastically altered by the rise of the living globules. Metabolism takes in substrates from the environment, transforms them into materials proper to the cells, and produces waste that is expelled into the environment. This situation may keep going

unrestrained as long as the environmental sources and sinks are not, respectively, exhausted and saturated. The organisms tend to spoil the environment they need to thrive in.

When the exhaustion/saturation limits are approached, the organisms are exposing themselves to new adaptive challenges that require other metabolic innovations. This means that evolutionary innovations become necessary components of the living process. The counterpart on the environmental side is that its degradation is inevitable and also a necessary character of the life process. Exploration of this side leads to the *Gaia* reasoning, which says that planets harboring living systems can be recognized through identification of unexpected compositions of their atmospheres.

Specificity starts with the genetic code The evolution of the endogenization process reached a peak in the development of specificities and precision, which guarantee speed and efficiency. These peaks were only possible with development of the genetic code. The code is the set of specific correspondences between triplets of nucleobases at a special site (anticodon) in the tRNA molecules and the amino acid that this tRNA carries at another special site. The correspondence is reached by the mutually adjusted sequence recognition between the tRNA and the enzyme that 'reads' the tRNA sequence and adds the correct amino acid to it.

With its nearly 'digital' resource (1 triplet → 1 amino acid), sequences of 'letters' could be enchained in polymers that would repetitively give origin to structures, catalysts, and transporters with guaranteed structure and function. It is tempting to propose that even the basic property of homochirality comes into biopolymers first through proteins, more precisely with Serine, the second amino acid encoded, since Glycine is nonchiral. Selection for the RNA constitution would also be related to the first module of encodings, Glycine being among the most frequent amino acids in RNA binding motifs. These two reasonings are monofactorial – protein-first, departing from the general coevolutionary message of the SRM, which is – more carefully stated – just RNP-first.

Memory in strings Having configured at least a few codes, those triplet sites could be arranged into ligated sequences that evolved high stability. Stability was obtained in (1) the 'static' sense of long duration, that is, resistance to degradation, by having (1a) the RNAs protected by the binding and coating by protective proteins, and then (1b) making copies of the RNA into DNA, which is chemically more stable. In (2) the more 'dynamic' sense, the stability of the sequence was obtained by choosing the efficient strategies of (2a) replication, utilizing the sequences as templates, and of (2b) correcting for replication mistakes, which is present in the DNA replication systems. These combined stabilities made the nucleic acids the ideal substrates for a memory function in polymer sequences called '*memory in strings'*.

The nucleic acid sequences are then called (A) genes, referring to their high stability and heritability due mostly to DNA, or (B) messenger RNAs (mRNAs), referring to the property of being translated into protein sequences. Other categories of RNAs (all are copied – transcribed – from DNA) serve the functions of regulators of the expression of the genes into the mRNAs, the tRNAs, and the ribosomal RNAs, all components of the machinery through which the mRNAs are translated via the tRNAs.

Cycles of repetitiveness The SRM proposes it has found the beginnings of the essentially circular mechanism: (a) tRNA anticodons are encoded by proteins to carry amino acids; (b) the anticodon complements, now in the codon function, are enchained into sequences of triplets in the nucleic acids genes and mRNAs; (c) these are translated into proteins with the help of the tRNAs, now in the decoding function, to make the protein chains; (d) the repetition of the cycle, starting from (a).

The living system is full of other types of cycles that will not be detailed here, for want of adequate space. Most important are the self-feeding or self-activating cycles that are typical of epigenetic memories.

Membranes Why does the SRM put the seeds for 'all' biological characters together in one and the same proto-RNP structure? Could they all be reduced to just the properties of the two informational biopolymers? Nothing is proposed here as certain, but the model seems to be plausible and is presented for evaluation. If judged adequate, it is open for improvements.

The 'spongy RNP globule' and the genetic code are mute with respect to membranes and lipids. Nonetheless some equivalent surface properties have been proposed to be accomplished by non-lipid constitutions, such as Sydney Fox's microspheres and Alexander Oparin's coacervates. In fact, natural membranes are never purely lipid, but lipoprotein in proportions ranging from 1 to 3 of each kind (25% to 75%) [18]. The surface of the globules could have started as protein-only, and lipid monomers would aggregate around some hydrophobic sites and start the lipoprotein membrane self-organization process. At the limit, vesicle formation by amino acids alone has been obtained in by Hyman Hartman's group, with units constituted of Aspartyl-$(Glycine)_{4\text{-}10}$, respectively, acidic head-hydrophobic tail [19].

Exploratory behaviors The restless, non-stop motility that characterizes living beings' protoplasm may find an analogy in the trembling and vibration of proteins in water, or in the variations in 3D conformations already cited in other contexts. Proteins at the globular surfaces presenting this behavior would be the seeds for the selection of the receptor and simple transporter functions.

Two preferential states among the variety of protein 3D conformations need to be selected for the final states. One has an extended or relaxed conformation, with sites exposed on the surface, open for interactions with external ligands. Upon interaction, the site gets closed, which leads to the acquisition of another state by the protein, where the site bearing the ligand is retracted to the inside of the globule. On meeting the internal milieu, the site opens again, the ligand is released, and the first conformation is reconstructed. This is the job of receptor cycling between two states: the open state is exposed outside and with longer duration, as if waiting/searching for substrates; the closed state moves inwards and quickly releases the ligand and goes out again. The material taken up is transferred to internal traps or to transformations along transduction and further processing routes.

This could be a reasonable route to go from the spongy RNP globule to the cell with receptors. Selection and adaptation acting upon protein pleomorphic states or changing 3D conformations would also bear analogies with the process of generating the exploratory behaviors of organisms as generic 'search and capture' procedures of cells.

Construction of the cellular flow system In short, the argument is that the protein synthesis sink mechanism has to be maintained healthy and, if not continually or perennially active, at least fully capable of resuming activity as soon as environmental conditions are adequate, in cases where it had to be temporarily suppressed due to harmful intercurrences. Simple examples of this kind occur in desiccation, freezing, lyophilization, and some cases of transforming cells into dormant spores and cysts. These cases introduce a third category in the doublet of being alive (metabolizing) or dead (incapable of having the lost metabolism restored), which is that of suspended life (not metabolizing but capable of restoration).

A first need is for diversification of protein structures and functions, to guarantee energy and amino acid sources in the upstream direction, and safe transport of products away from the synthesis sites in the downstream direction. There must be no clogging, blockades, and accumulations along the flow routes. The protein-synthesis system works as a substrate-stimulated ratchet, not presenting a drive-forward mechanism itself. Diversification of proteins depends mainly on gene duplication, mobility, and horizontal transfers, but may incorporate also epigenetic influences. All these mechanisms may be grouped in the concept of plasticity, genomic or phenotypic.

The living mechanism incorporates reversion of the polymerization direction only for generating monomers at nutritional salvage. Sensitivity to saturation is one of the regulatory processes of the polymerization activity. Mechanical saturation is avoided through control of cell volumes and shapes, avoiding the effects of overcrowding through skeletal features such as the microtubules and filaments of eukaryotes, and the cell walls of plants, fungi, and prokaryotes. If saturation does not work by itself, it triggers the activation of repressors. Saturation may not need to be general, but may be restricted to some specific kinds of processes that developed the role of critical sensors for control.

Lincoln, Maxwell Diffusion inside cells is largely tortuous and limited by compartments. Much of the water is not available for free flow but adhered to and organized around biopolymers, so that transport prevails over diffusion and flow becomes more controlled than free. The most complex eukaryotic cells may develop structures devoted specifically to flow mechanics, polarity, etc. The simpler prokaryotes may have entry and disposal sites in large numbers and distributed on the surfaces.

In spite of the tortuosity, internal structures have to be at least flow-friendly, if not flow-promoting, which allows that 'the flow organizes the structures, and the latter help that same process'. Paraphrasing Abraham Lincoln's Gettysburg address on democracy, the cell is made with the flow, for the flow, by the flow… Products of metabolic pathways have to be moved away from the synthesis sites or to be modified anew, so that the pathway acquires directionality and does not revert to its old state due to mass-action effects. This is also the crucial structure in the famous "Maxwell's Demon" that generates order from probabilistic distributions: enzymes give speed and specificity which are combined with the flow and gradient directionality.

Accumulated mistakes and lesions The fragility of the 'spongy globule' obtains continuous activity of the protein synthesis sink with spontaneity due to its unsaturated state, not reaching overcrowding. The aggregation between components and development of membranes is rudimentary, and is only on the way to later guarantee resistance to fragmentation and to invasion by water. This is documented by Fox's microspheres and Oparin's coacervates photographs of gemmulation similar to the oocyte-polar body or the budding yeast mother-daughter cell associations. In fully developed cells this mechanism of losing chunks of protoplasm, now called the shedding of vesicles or exosomes, became regulated but reaches the same stimulatory effect.

This process was profited from in other ways, having its functionality enriched. Getting rid of waste from carbon of sugars and lipids is well resolved with its transformation into gases that are also highly soluble and evaporate easily. Nitrogenous compounds remain a serious problem. Its derivatives are toxic (ammonia, insoluble) urate, or water-requiring and pollutant (urea). Some amino acids and some proteins are not well reabsorbed in the kidneys and are disposed of as such. Some intracellular proteins are most difficult to degrade and cannot be recycled through catabolic processes such as the proteasomal and the lysosomal-autophagic processes. Degradation intermediates that cannot be further processed may be toxic to cellular organization, especially via the exposition of the unprotected internal hydrophobic protein cores. After these damages, they reach a final state as aggregates, tangles, and amyloid grains and plaques. The only solution left is vesicle shedding.

The SRM introduced the consideration of 3D protein construction rules (folding) among the tests and components of its structure. Initial empirical data came from Varshavsky's N-end Rule. This describes which amino acids contribute to protein resistance to degradation when placed at their N-ends. It contributed decisively to the configuration of the genetic code as a circular structure where the initiation and termination codes are the last to form and are dictated one by the other, producing an 'informational closure' that is also material. The N-end Rule is the mother of a large research line on cotranslational protein folding, whose rules describe the acquisition of 3D conformations that prevent the damage-prone formation

of tangles. In this way, a connection is reached between the SRM and the processes of protein metabolic stability and prevention of aggregation damages.

Good and bad Exosome multifunctionality includes, through the loss of biomass, the beneficial effect of the protein synthesis sink stimulation. Additional benefits are the substitution of aged material lost with renovated structures and functions. The stimulatory and regenerative mechanism is analogous to that obtained from tree pruning.

Medical application of this knowledge is nowadays twofold. Extracellular vesicles and exosomes are also seen as communication tools that cells utilize for transport of macromolecules inside multicellular bodies, and can be utilized by attendants as 'liquid biopsies' for laboratorial diagnosis. Intracellular accumulation of protein tangles is seen as a possible causative agent of various diseases, including Alzheimer's. Especially when neurons are involved, there is always the question of whether the accumulated material is a consequence of the aged post-mitotic state, the disease being caused by some other mechanism, or whether the accumulated material is the cause or a co-participant among the causes of the disease.

Reproduction The physiological rationale [13] is more easily extended when data on bacterial reproduction are taken into account. Cells that reach large sizes (such as in the G2 stage of the eukaryotic cell cycle) run the risk of having protein synthesis reduced/inhibited due to mass-action or saturation-induced repression. In these situations, exosome extrusion became regular and obligatory, in order to release them from inhibition in case they needed to continue to be healthy and productive. This is equivalent to the cytoplasmic fission in cell reproduction. The latter was reached when sets of genetic memories – genomes – were added to the chunks of cytoplasm being eliminated, these becoming daughter cells. The idea is that the original function of the first phase in the reproduction process – cytoplasm fission – is that of regenerating the protein synthesis activity, while the second function was that of rescuing the cytoplasm portions from loss by becoming daughter cells.

Bacterial L-forms, peeled off from the walls, have a more fragile membrane, and the exosome formation is easily observed. In some of the exosomes, genomes are included, leading to the proposition of this being a primitive form of reproduction [2,12,14]. In other studies, it is seen that cell reproduction may be asymmetric with respect to the inclusion in only one of the daughter cells of an 'inclusion body' which contains the tangles of damaged and undigested protein. This is a nice way of producing healthier lineages, free from the tangles, separated from the less healthy lineages [17]. The segregation is simplified due to the tendency of all damaged proteins to get entangled together in one inclusion body.

Behaviors focus on the extremities of the flow system Reproduction having been reached, and its origins and functions understood, 'informational closure' receives consistency. Real evolutionary populations are formed when the Darwinian process reached completion. Whenever reproduction is active, the protein synthesis activity may be never-ending, subjected to the evolutionary openness.

The main foci of organismal open behaviors, the most apparent 'vital force' manifestations, are found in the extremes of the process: (α) nutrition, that affects protein synthesis sink, and (Ω) reproduction, that pulls the sink forward and keeps it active nonstop. Intermediate mechanisms are internal organic ones, which may go unnoticed by external observers as they are mostly hidden to the organic senses and the conscious feelings of individuals.

Acknowledgments To all of my colleagues in the Self-Organization Group at UNICAMP, who have accompanied me in a journey of almost 30 years, and to the many teachers and collaborators in the busy but happy life we lead.

Conflicts of interest I declare no conflicts of interest between this publication and my other activities.

References

1. Chistoserdova L 2011 Modularity of methylotrophy, revisited. Environ. Microbiol. 13, 2603–2622

2. Errington J 2013 L-form bacteria, cell walls and the origins of life. Open Biol 3, 120143

3. Farias ST, Rêgo TG, José MV 2016 tRNA core hypothesis for the transition from the RNA world to the ribonucleoprotein world. Life 6, e15

4. Florio R, Salvo ML, Vivoli M, Contestabile R 2011 Serine hydroxymethyltransferase: A model enzyme for mechanistic, structural, and evolutionary studies. Biochim. Biophys. Acta 1814, 1489–1496

5. Gayon J 2010 Defining life: synthesis and conclusions. Orig Life Evol Biosph 40, 231–244. Also: Why is it so difficult to define what is life and what are living beings. 5 February 2017 http://www.bbc.com/portuguese/vert-earth-38800106

6. Goudry M, Saugnet L, Belin P, Thai R, Amoureux R, Tellier C, Tuphile K, Jacqet M, Braud S, Courçon M et al. 2009 Cyclodipeptide synthases are a family of tRNA-dependent peptide bond-forming enzymes. Nat. Chem. Biol. 5, 414–420

7. Guimarães RC 2014 Essentials in the Life Process Indicated by the Self-Referential Genetic Code. Orig Life Evol Biosph 44, 269–277 doi 10.1007/s11084-014-9380-7

8. Guimarães RC 2015 Emergence of information patterns: In the quantum and biochemical realms. Quantum Biosyst. 6, 148–159

9. Guimarães RC 2017 Self-Referential Encoding on Modules of Anticodon Pairs—Roots of the Biological Flow System. Life 7, 16 doi 10.3390/life7020016. *This review contains a large collection of references*.

10. Ikehara K 2009 Pseudo-Replication of [GADV]-Proteins and Origin of Life. Int. J. Mol. Sci. 10, 1525–1537

11. José MV, Zamudio GS, Palacios-Pérez M, Bobadilla JR, de Farias ST 2015 Symmetrical and thermodynamic properties of phenotypic graphs of amino acids encoded by the primeval RNY code. Orig. Life Evol. Biosph. doi 10.1007/s11084-015-9427-4

12. Leaver M, Domínguez-Cuevas P, Coxhead JM, Daniel RA, Errington J 2009 Life without a wall or division machine in Bacillus subtilis. Nature 457, 849–855

13. MacDonald IA, Kuehn MJ 2013 Stress-induced outer membrane vesicle production by Pseudomonas aeruginosa. J Bacteriol 195, 2971–2981

14. Mercier R, Kawai Y, Errington J 2013 Excess membrane synthesis drives a primitive mode of cell proliferation. Cell 152, 997–1007

15. Mocibob M, Ivic N, Bilokapic S, Maier T, Luic M, Ban N, Durasevic IW 2010 Homologs of aminoacyl-tRNA synthetases acylate carrier proteins and provide a link between ribosomal and nonribosomal peptide synthesis. Proc. Natl. Acad. Sci. USA 107, 14585–14590

16. Moutiez M, Schmidt E, Seguin J, Thai R, Favry E, Belin P, Mechulan Y, Goudry M 2014 Unravelling the mechanism of non-ribosomal peptide synthesis by cyclopeptide synthases. Nat. Commun. 5, 5141–5146

17. Nyström T 2011 Spatial protein quality control and the evolution of lineage-specific ageing. Philos Trans R Soc London B 366, 71–75

18. Rizzotti M 2000 Early Evolution. Birkhäuser, Basel Switzerland

19. Santoso S, Hwang W, Hartman H, Zhang S 2002 Self-assembly of surfactant-like peptides with variable glycine tails to form nanotubes and nanovesicles. Nano Letters 2, 687–691 doi 10.1021/nl025563i

20. Sobolevsky, Y.; Guimarães, R.C.; Trifonov, E.N. 2013 Towards functional repertoire of the earliest proteins. J. Biomol. Struct. Dyn. 31, 1293–1300

21. Trifonov EN 2004 The triplet code from first principles. J. Biomol. Struct. Dyn. 22, 1–11

22. Weiss MC, Neukirchen S, Roettger M, Mrujavac N, Sathi SN, Martin WF, Sousa FL 2016 Reply to 'Is LUCA a thermophilic progenote?'. Nat. Microbiol. 1, 16230