 doi:10.20944/preprints201811.0620.v1

Communication

Life emerged as the "protein/metabolism-first" theory expects

Kenji Ikehara[1, 2, 3,*]

[1] G&L Kyosei Institute, Koharu Bld. 202, Hokkeji 153-4, Nara, Nara 630-8001, Japan. ; ikehara@cc.nara-wu.ac.jp

[2] The International Institute for Advanced Studies of Japan, Kizugawadai 9-3, Kizugawa, Kyoto 619-0225, Japan.

[3] Nara Women's University, Emeritus professor

* Correspondence: ikehara@cc.nara-wu.ac.jp; Tel.: +81-774-73-4478

Abstract: The origin of life has not been solved as yet, in spit of the time passage more than thirty years from publication of RNA world hypothesis by W. Gilbert (1986), which is based on the "gene/replicator--first" theory. On the contrary, I have proposed [GADV]-protein world hypothesis (GADV hypothesis), assuming that life emerged from [GADV]-protein world, which is grounded on the "protein/metabolism-first" theory. However, two weak points of protein world hypothesis, (i) protein cannot be produced without gene, and (ii) protein cannot be self-replicated, have been frequently pointed out by supporters of RNA world hypothesis. Then, I examined whether the two weak points could be overcome by GADV hypothesis or not. From the results, it was confirmed that (i) [GADV]-protein could be pseudo-replicated in the absence of gene owing to protein 0th-order structure or [GADV]-amino acids, and (ii) the replication ability is not always required from the beginning but it is sufficient to acquire it at some time point until the emergence of life. Thus, it was concluded that life emerged as [GADV]-protein world hypothesis, which is grounded on the "protein/metabolism-first" theory, expects.

Keywords: gene-first 1; protein-first 2; replicator-first 3; metabolism-first 4; origin of life 5; [GADV]-protein world hypothesis 6; GADV hypothesis 7; RNA world hypothesis 8

---

1. Introduction

Problem about how life emerged on the primitive Earth, or about which is correct, the "gene/replicator-first" or "protein/metabolism-first" theory, was disputed for many years. The disputes continued until about 10 years ago [1-3]. After the disputes, actually since RNA world hypothesis was proposed by Gilbert [4] upon discoveries of ribozyme or catalytic RNA [5, 6], life-origin studies have been mainly carried out from standpoint of the "gene/replicator-first". The reason is because it was considered by many researchers that so-called the "chicken-egg relationship" observed between gene and protein could be explained and the problem about the origin of life could be settled by the RNA world hypothesis. However, the fact, that the riddle on the origin of life is unsolved still now despite strenuous efforts by many researchers, would indicate that the problem has not come to the conclusion [7]. For example, it has been pointed out that studies from "ground zero" will be necessary to solve the riddle on origin of life judging from the present situation of the studies on the origin of life [8, 9].

On the other hand, I have proposed [GADV]-protein world hypothesis (GADV hypothesis), suggesting that life emerged from[GADV]-protein world [7, 10, 11]. [GADV] means four amino acids, Gly [G], Ala [A], Asp [D] and Val [V]. GADV hypothesis is, of course, an idea discussing the origin of life from the standpoint of "protein/metabolism-first".

I will show in this communication that life emerged on the primitive Earth as the "protein/metabolism-first" theory expects.

2. Did the first life emerge as expected by the "protein/metabolism first" theory?

2.1. Two weak points in "protein/metabolism-first" theory or protein world hypothesis

It is widely considered that there are two weak points in the "protein/metabolism-first" theory, as described below (Table 1).

(i) Protein synthesis is impossible in the absence of gene.

Of course, only random reactions should occur on the primitive Earth. In addition, one amino acid sequence generally corresponds to one structure of a protein. Therefore, extant protein like as precision polymer machine could not be produced in the absence of gene or by random joining of amino acids. The assertion is totally correct, since one amino acid sequence must be selected out from extraordinarily wide sequence space as $20^{100} = \sim 10^{130}$, to produce the protein (Fig. 1A) [12].

(ii) Protein replication is also impossible.

Replication is indispensable for the emergence of life. Nevertheless, protein cannot be replicated like as nucleic acids, DNA and RNA, because there is not one to one correspondence between two amino acids like as base pair of DNA and RNA.

Taken the two impossibilities of protein into consideration, it could not be expected that life emerged as the "protein/metabolism-first" theory conjectures. Therefore, it would be natural to consider that life must emerge as expected by the "gene/replicator-first" theory.

Table 1. Two weak points of protein world hypothesis on the origin of life.

Hypothesis	Mature protein synthesis [1]	Protein self-replication
Protein world	Impossible without gene	Impossible

[1] Mature protein means a complete one, which has matured to such as an extant protein like a precision polymer machine.

The reason, why the two assertions described above were led, would be because two important points have been overlooked. (i) It is unnecessary to produce protein as one like a precision polymer machine from the beginning, and it should be sufficient to produce immature water-soluble globular [GADV]-proteins with slightly more flexible structure than extant mature proteins (Fig. 1B). [GADV]-amino acids, one of protein 0^{th}-order structures, would make it possible to produce the immature protein in the absence of gene on the primitive Earth [13]. The immature protein could be evolved to mature protein, after double-stranded gene was formed (Fig. 1C). That is, it is unnecessary to assume the "gene-first" for the emergence of life. (ii) It is also unnecessary to equip with the ability of self-replication for the emergence of life from the beginning, although the ability must be, surely, indispensable for life to live. Not the case, it is sufficient if the ability could be acquired at some time point during chemical evolution from inanimate chemical compounds to the emergence of life (Fig. 2). In other words, it is unnecessary to assume the replicator-first for the emergence of life.

These points will be discussed in the following Section 2.2. in more detail.

2. 2. [GADV]-protein world hypothesis grounded on the "protein/metabolism-first" theory

Here, I will show that the two weak points described in Section 2.1 (Table 1) can be overcome by [GADV]-protein world hypothesis, which I have proposed [7, 10, 11]. Outline of evolutionary

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2018 doi:10.20944/preprints201811.0620.v1

process to the emergence of life, which was expected by the GADV hypothesis, is described below [14].

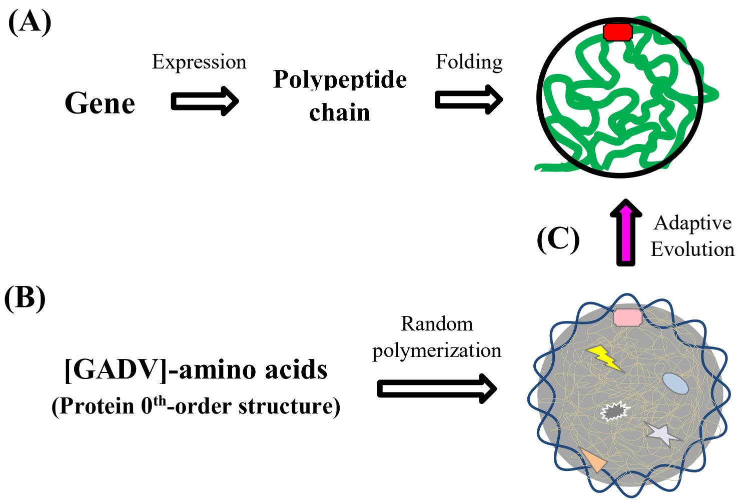

Fig. 1. **(A)** Polypeptide chain (green thick curve) synthesized by gene expression is generally folded into one tertiary structure (black circle), which has one active center catalyzing a chemical reaction against one substrate (red ellipse). **(B)** Random polymerization of [GADV]-amino acids in a protein 0^{th}-order structure or, for example, in a pool containing [GADV]-amino acids at roughly equal amount, could produce water-soluble globular protein with flexible structure. A number of active centers for various organic compounds should appear on surface of the protein, owing to the flexible protein structure. Various forms of faintly colored marks represent various organic molecules, which could be weakly catalyzed by the immature protein. Pinkish ellipse indicates weak active center before adaptive evolution. **(C)** If one catalytic activity, which is necessary to live, could be found on the flexible protein, the immature protein could gradually evolve to form mature protein like a precision polymer machine after invention of double-stranded gene.

Stage 0: *Before formation of [GADV]-protein world*

Various kinds of organic compounds, as amino acids, amines, organic acids and so on, were synthesized through lightning in primitive atmosphere [15], at hydrothermal vents in deep sea [16, 17] and *etc.* and accumulated on the primitive Earth in the absence of proteinaceous catalyst. The organic compounds could be also delivered to the early Earth by meteorites, asteroids, comets, and interplanetary dust particles [18].

Stage 1: *Formation of [GADV]-protein world*

[GADV]-protein, actually aggregate of [GADV]-peptides, could be synthesized such as by repeated heat-drying process in depressions on rocks (protein-first) [19]. Successively, [GADV]-protein world was formed upon accumulation of [GADV]-proteins produced by pseudo-replication (Fig. 2 (A)) [20]. Pseudo-replication is the synthesis of water-soluble globular [GADV]-proteins with similar amino acid composition but not the same amino acid sequence through direct random joining of [GADV]-amino acids. Thus, the most primitive metabolic system could be formed in the [GADV]-protein world (metabolism-early) (Fig. 2 (A)).

Stage 2: *Establishment of GNC genetic code*

Oligonucleotide synthesis was carried out by [GADV]-proteins, which led to establishment of GNC primeval genetic code through stereospecific interaction between one of [GADV]-amino acids and oligonucleotide containing GNC anticodon, which corresponds to the respective [GADV]-amino acids. The GNC-containing oligonucleotide became the most primitive tRNA [10, 11, 14]. Side-by-side dimer formation between the two primeval tRNAs could stimulate [GADV]-protein synthesis because of adjoining effect of two amino acids through the two complexes [14]. The complexes could also form another type of dimer, one complex opposites to

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2018 doi:10.20944/preprints201811.0620.v1

another complex through base pairing between two complementary GNC anticodons in the complexes (Fig. 2 (B)) [14, 21]. Water-soluble globular protein containing only [GADV]-amino acids could be produced under the first GNC genetic code for the first time. Protein 0th-order structure or [GADV]-amino acids made it possible to synthesize water-soluble globular [GADV]-protein in this case too.

Stage 3: *Single-stranded (GNC)$_n$ gene formation*

Single-stranded (GNC)$_n$ RNA gene could be formed by random joining of GNC anticodons in the complexes (gene-late). The protein 0th-order structure, [GADV]-amino acids, also led to synthesize water-soluble globular [GADV]-protein with the first single-stranded RNA gene, because primeval GNC genetic code encoded [GADV]-amino acids.

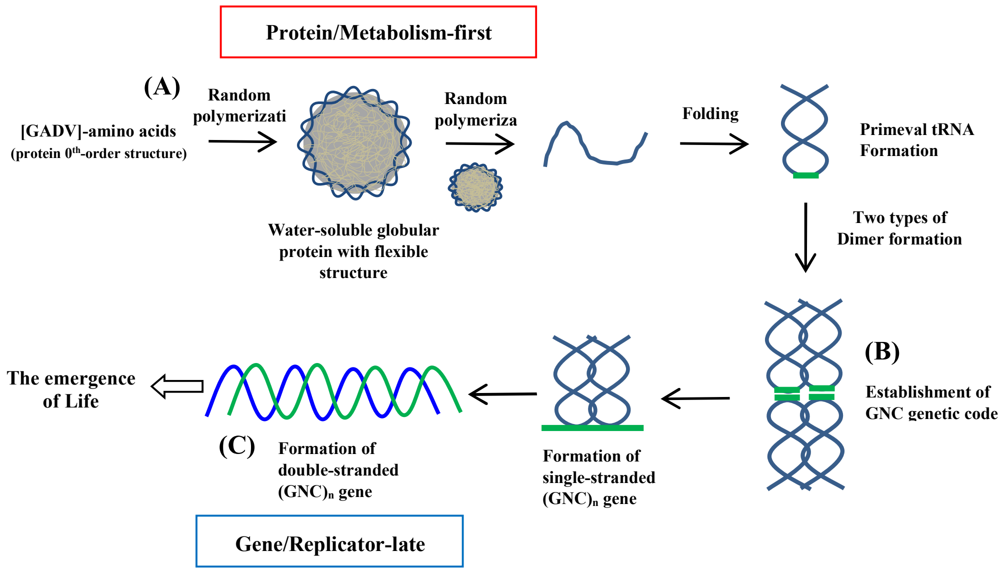

Figure 2. Formation process of the first primeval genetic system composed of gene, genetic code and protein: The formation of the system began with (A) synthesis of immature water-soluble globular [GADV]-protein with slightly flexible structure by random joining of [GADV]-amino acids, which was followed by (B) establishment of GNC genetic code and (C) creation of double-stranded (GNC)$_n$ gene. It is clearly indicated that formation of the system was achieved according to the "protein/metabolism-first theory". Thick short and long green lines indicate anticodon and single-stranded gene, respectively.

Stage 4: *Double-stranded (GNC)$_n$ gene formation*

Double-stranded (GNC)$_n$ RNA gene was produced through complementary strand synthesis of the first single-stranded (GNC)$_n$ gene (Fig. 2 (C)). Immature, flexible [GADV]-protein with a weak catalytic activity could first evolve to a mature, rigid protein with a high catalytic activity through introduction of required base substitutions onto the sense sequence, after formation of double-stranded (GNC)$_n$ gene (Fig. 1 (C)). Propagation of genetic information to progeny also became possible through gene replication at the last stage.

Therefore, it can be supposed that the process to the emergence of life progressed in order of [GADV]-protein synthesis in the absence of gene (stage 1), development of primitive metabolism using [GADV]-proteins (stage 1), establishment of primeval GNC genetic code (stage 2), formation of the first single-stranded (GNC)$_n$ gene (stage 3) and creation of replicator or double-stranded (GNC)$_n$ gene (stage 4). After the creation of a number of double-stranded genes necessary to live, the first life emerged. Note that the processes based on the GADV hypothesis are clearly followed as the "protein/metabolism-first" theory expects (Fig. 2 (A), Table 2). On the contrary, double-stranded gene or replicator was formed at the last stage 4 or just before the emergence of life (Fig. 2 (C), Table 2).

Table 2. The stage of protein/metabolism or gene/replicator formation, which is assumed by [GADV]-protein world hypothesis

Hypothesis	Formation of protein/metabolism	Formation of gene/replicator
[GADV]-protein world	At the first stage	At the last stage

2.3. Conditions of molecules for triggering off the emerging of life

Molecules leading to the emergence of life must satisfy the following conditions.

1. Organic molecules used for the emergence of life must be synthesized by physical, physicochemical and chemical reactions on the primitive Earth.

For satisfying the first condition, organic molecules should be composed of atoms, of which atomic number is as small as possible. In addition, the number of atoms contained in the organic molecules should be as few as possible.

2. Organic molecules must be useful for leading to the emergence of life.

For satisfying the second condition, chemical reactions, which are necessary to lead to the emergence of life, should be catalyzed by organic molecules, in which functional groups with positive and negative charges are contained.

3. Organic molecules must be used, which could trigger off formation of the first fundamental life system composed of gene, genetic code and protein.

This means that the organic molecules should be one of components of RNA (DNA), tRNA or protein.

2.4. Grounds supporting GADV hypothesis based on "protein/metabolism-first" theory

Here, I will discuss whether [GADV]-protein and [GADV]-amino acids can satisfy the above three conditions for triggering off the emergence of life or not, in order to confirm adequacy of GADV hypothesis..

2.4.1. [GADV]-amino acids could be easily synthesized on the primitive Earth

The first gene and tRNA realizing genetic code are polymers of nucleotides composed of H, C, N, O and P. The number of atoms of four nucleotides (AMP, UMP, CMP and GMP) is within the range from 34 to 37. On the contrary, [GADV]-protein is a polymer of [GADV]-amino acids composed of H, C, N and O. The number of [GADV]-amino acids is ranging from 10 to 19. Therefore, [GADV]-amino acids should be more easily synthesized on the primitive Earth than nucleotides, because the amino acids are much simpler than nucleotides (Condition 1). In addition, the amino acids have positive charge as amino group and negative charge as carboxyl group in the molecules. Therefore, peptide bond could be easily formed between the two amino acids in water owing to positive and negative charges, indicating that [GADV]-amino acids satisfy the condition 2. Needless to say, [GADV]-amino acids, which are components of protein, also satisfy the third condition, which is necessary to lead to the emergence of life.

2.4.2. [GADV]-protein, which was synthesized by random joining of [GADV]-amino acids, could be folded into water-soluble globular structure

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2018 doi:10.20944/preprints201811.0620.v1

The reason, why [GADV]-amino acids were used for constituents of the first protein, would be because the amino acids are rather smaller molecules in 20 natural amino acids. Certainly, Gly carrying H atom as side chain is the smallest amino acid and Ala having methyl group is the second smallest amino acid in all amino acids, respectively. However, it is obvious that the amino acids were not always selected out from all amino acids in order from the smallest amino acid, because the numbers of atoms of serine and cystein are smaller than that of valine. Furthermore, nonnatural 2-amino butylate, which has ethyl group as side chain, is also simpler than valine. The reason, why those amino acids were not used for the first protein, is because serine is turn/coil forming amino acid and both cystein and 2-amino butylate without bulky side chain at β-carbon atom are α-helix forming amino acids. In contrast, valine carrying two methyl groups at β-carbon atom is β-sheet forming amino acid. Therefore, one of four conditions for formation of water-soluble globular protein or β-sheet formability, is not satisfied, if serine, cystein or 2-amino butylate was used for the first protein instead of valine. Gly, Ala and Val are the smallest turn/coil, α-helix and β-sheet forming amino acids in all amino acids, respectively. Asp is hydrophilic amino acid carrying carboxyl group and Val is hydrophobic amino acid. In other words, [GADV]-amino acids are a combination of the smallest amino acids, which can satisfy the four conditions for water-soluble globular protein formation [22].

2.4.3. [GADV]-protein was the fittest polymer leading to the emergence of life

Next, I will explain that [GADV]-protein and its components, [GADV]-amino acids, are the fittest organic molecules for leading to the emergence of life, taken experimental results and the results of computer analysis of microbial genes and proteins, which were obtained thus far, into consideration.

1. The results, which were obtained by experiments carried out with Miller's type electric discharge and apparatus simulating hydrothermal vent and by chemical analysis of meteorites, are well summarized in the papers published by Higgs [23] and van der Gulik *et al.* [24], and it is confirmed that [GADV]-amino acids could be easily synthesized and accumulated on the primitive Earth.

2. Peptide bond could be also easily formed between [GADV]-amino acids, for example, by repeated heat-drying process in a depression of rock on seaside of the primitive Earth, because of electric attraction acting between positive charge on amino group and negative charge on carboxyl group of [GADV]-amino acids [19]. Similar results were also reported in 1988 by Yanagawa *et al.* [25].

3. Luisi has provided the results in his book showing that even simple peptides as Gly-Gly can catalyze peptide bond hydrolysis [9]. The results indicate that [GADV]-proteins, which were produced even by random joining of [GADV]-amino acids, also could have some catalytic activities. In fact, it has been experimentally shown that [GADV]-protein as an aggregate of [GADV]-oligopeptides can catalyze hydrolytic reactions of peptide bond in natural protein, bovine serum albumin, and of phosphodiester bond in tRNA [19].

4. It is expected that even polypeptide chain synthesized by random joining of [GADV]-amino acids could be folded into water-soluble globular structure, because [GADV]-protein can satisfy the four conditions for formation of water-soluble globular protein, which were obtained by computer analysis of extant water-soluble globular proteins from seven microorganisms [26].

5. It is supposed that structure of [GADV]-protein produced by random joining of [GADV]-amino acids were more flexible than that of extant protein, because turn/coil forming Gly should accumulate on the primitive Earth at a larger amount than β-sheet forming Val and, therefore, Gly would be incorporated into the [GADV]-protein more than Val. Use of a large amount of Gly would inhibit secondary structure formation and weaken hydrophobic core of the protein [22]. Furthermore, one immature and flexible [GADV]-protein could make it possible to play an important role in exhibiting catalytic activities against various substrates (Fig. 1 (B)). The reason is because the protein could easily adapt surface amino acid residues to many kinds of organic

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 November 2018 doi:10.20944/preprints201811.0620.v1

compounds owing to the flexible protein structure. On the contrary, RNA, of which structure is mainly folded with intramolecular hydrogen bonds, would be too rigid to adapt against substrate, suggesting that it is difficult for RNA synthesized by random joining of nucleotids to express catalytic activity against various organic compounds.

6. In addition, [GADV]-protein could exhibit catalytic activities of nucleotide metabolism, such as RNA polymerase and kinase activities [24]. Those activities could contribute to produce the first primeval tRNA, which should trigger off formation of the first single-stranded RNA gene (Fig. 2 (B)).

Thus, unlike RNA, [GADV]-protein could be easily synthesized through physical, physicochemical and chemical reactions and could exhibit various catalytic activities leading to the emergence of life through formation of the first double-stranded RNA gene (Fig. 2 (C)).

3. Discussion

It was confirmed that [GADV]-protein world hypothesis can explain evolutionary process to the emergence of life without any large contradiction, in spite that the idea is grounded on the "protein/metabolism-first" theory, in which two weak points, which it might be impossible to overcome, are included (see Section 2. 1).

On the contrary, RNA world hypothesis is an idea, which is grounded on the "gene/replicator-first" theory. If that is correct, first of all, it causes a problem what the gene is in the gene-early theory. Does the gene encode an enzyme as like a precision polymer machine? However, such a gene could never be created first, because the first gene must be formed by random polymerization of mononucleotides, if the gene were made of RNA. Extraordinary large sequence diversity (the diversity of RNA encoding even a small protein composed of only 100 amino acids reaches $(4^3)^{100}$ = about 10^{180}) should make formation of the first gene quite difficult or actually impossible.

Metaphorically speaking, it is an idea that a masterpiece could appear during writing alphabets randomly. Even the probability that a simple phrase, "origin of life", appears during random writing, reaches $1/27^{14}$ = about $1/10^{20}$, if space between two words is included as like one letter in twenty-six alphabets. This means that even a short novel never appear during random writing of alphabets. As a matter of course, an excellent piece of music, similarly, never be composed by random arrangement of notes on a sheet music. Letters and notes were not generated independently of concrete objects and of imagination in mind but invented to express objects and melodies and to record sentences and musics through joining words and melodies, respectively. Similarly, a gene or genetic information was created to express amino acid sequence of a protein one day in the past through codons of concrete objects or amino acids, and is used to reproduce a protein in organisms on the Earth. Thus, information never appear independently of entities in every world of literature, music, life and so on. It is concluded here that the "gene-first" theory on the origin of life as RNA world hypothesis never be realized.

In addition, replicator RNA as a chemical material would not acquire any genetic information even after repeated base substitutions, because the self-replicated RNA produced independently of protein would evolve only to improve the ability of self-replication. Furthermore, the replicator or self-replicated RNA would withdraw into the RNA world and the RNA world could not develop to RNA-protein world, because the replicator RNA with no relation to protein would not require genetic information for protein synthesis. In fact, evidence showing that RNA world developed to RNA-protein world has not been provided until now, in spite that more than thirty years already passed from proposition of the RNA world hypothesis [4].

As a matter of course, the extant life system is composed of both gene as a replicator and protein with catalytic function, which is necessary to metabolize organic compounds in cell. As described above, [GADV]-protein world was first formed by pseudo-replication of [GADV]-protein (Fig. 1 (B); Fig. 2 (A)), which was become possible owing to the protein 0^{th}-order structure or

[GADV]-amino acids [20]. However, the [GADV]-proteins could not withdraw in the protein world, because protein itself cannot be replicated and evolved. Therefore, [GADV]-protein asked RNA to assist reproduction of the protein. The necessity of RNA led to establishment of GNC primeval genetic code (Fig. 2 (B)) and to formation of the first double-stranded RNA gene as the result (Fig. 2 (C)). This is the reason, why the genetic system was formed in order of protein and RNA, but not of RNA and protein and the first gene or replicator must be produced after [GADV]-protein formation. Thus, the so-called "chicken-egg relationship" between gene and protein was established at a time just before the emergence of life. Thereafter, the relationship has been maintained without interruption until now. Inversely, the relationship has forced not to get separated from each other.

Then, why has the "gene/replicator-first" theory, which could not be realized during the process to the emergence of life, been accepted by many researchers up to the present time? The reason would be because they do not sufficiently recognize the significance of GADV hypothesis based on protein 0th-order structure or a specific amino acid composition of [GADV]-amino acids. Generally speaking, it would be unavoidable to reach to a wrong conclusion, whenever it is considered in a situation as overlooking a quite important concept, like as the case of the "gene/replicator-first" theory.

Acknowledgments: I am very grateful to Dr. Tadashi Oishi (G&L Kyosei Institute, Emeritus professor of Nara Women's University) for encouragement throughout our research on the origin and evolution of the fundamental life system .

Conflicts of Interest: The author declares no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Vasas, V.; Szathmary, E.; Santos, M. Lack of evolvability in self-sustaining autocatalytic networks constraints metabolism-first scenarios for the origin of life. *Proc. Natl. Acad. Sci.* **2010**, *107*, 1470-1475.
2. Pross, A. Causation and the Origin of Life. Metabolism or Replication First? *Orig. Life Evol. Biosph.* **2004**, *34*, 307-321.
3. Shapiro, R. A replicator was not involved in the origin of life. *IUBMB Life* **2000,** *49*, 173-176.
4. Gilbert, W. The RNA world, *Nature* **1986**, *319*, 618.
5. Kruger, K.; Grabowski, P.J.; Zaug, A.J.; Sands, J.; Gottschling, D.E.; Cech, T.R. Self-splicing RNA: autoexcision and autocyclization of ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* **1982**, *31*, 147-157.
6. Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. The RNA moiety of ribonuclease P is catalytic subunit of the enzyme, *Cell* **1983**, *35*, 849-857.
7. Ikehara, K. Evolutionary Steps in the Emergence of Life Deduced from the Bottom-Up Approach and GADV Hypothesis (Top-Down Approach). *Life* **2016**, *6*, 6. doi:org/10.3390/life6010006
8. Luisi, P.L. A new start from ground zero? *Orig. Life Evol. Biosph.* **2014**, *44*, 303–306.
9. Luisi, P.L. "*The emergence of life-from chemical origins to synthetic biology-* 2nd ed.", Cambridge Univ. Press, Cambridge, United Kingdom, 2016.
10. Ikehara, K. Origins of gene, genetic code, protein and life: Comprehensive view of life system from a GNC-SNS primitive genetic code hypothesis. *J. Biosci.* **2002**, *27*, 165–186.
11. Ikehara, K. Possible steps to the emergence of life: The [GADV]-protein world hypothesis. *Chem. Rec*. **2005**, *5*, 107–118.
12. Dill, K.A. Dominant forces in protein folding. *Biochemistry*, **1990**, *29*, 7133–7155.
13. Ikehara, K. Protein ordered sequences are formed by random joining of amino acids in protein 0th-order structure, followed by evolutionary process. *Orig. Life Evol. Biosph.* **2014**, *44*, 279–281.
14. Ikehara, K. Life Emerged from [GADV]-Protein World, but Not from RNA World!?. *Preprints* **2017**, 2017120170 (doi: 10.20944/preprints201712.0170.v1)
15. Miller, S.L.; Orgel, L.E. In *The Origins of Life on the Earth*, Prentice-Hall: Englewood Cliffs, NJ, USA, 1974.
16. Imai, E.; Honda, H.; Hatori, K.; Brack, A.; Matsuno, K. Elongation of oligopeptides in a simulated submarine hydrothermal system. *Science* **1999**, *283*, 831–833.

17. Holm, N.G.; Andersson, E. Hydrothermal simulation experiments as a tool for studies of the origin of life on Earth and other terrestrial planets: A review. *Astrobiology* **2005**, *5*, 444–460.
18. Botta, O.; Glavin, D.P.; Kminek, G.; Bada, J.L. Relative amino acid concentrations as a signature for parent body processes of carbonaceous chondrites. *Orig. Life Evol. Biosph.* **2002**, *32*, 143–163.
19. Oba, T.; Fukushima, J.; Maruyama, M.; Iwamoto, R.; Ikehara, K. Catalytic activities of [GADV]-peptiders. Formation and establishment of [GADV]-protein world for the emergence of life. *Orig. Life Evol. Biosph*. **2005**, *35*, 447–460.
20. Ikehara, K. Pseudo-replication of [GADV]-proteins and origin of life. *Int. J. Mol. Sci*. **2009**, *10*, 1525–1537.
21. Guimarães, R.C. Self-Referential Encoding on Modules of Anticodon Pairs—Roots of the Biological Flow System. *Life* **2017**, 7, 16; doi:10.3390/life7020016
22. Ikehara K. In *GADV Hypothesis on the Origin of Life -Life emerged in this way!?-*, LAMBERT Academic Publishing: Saarbrucken, Germany. 2016.
23. Higgs, P.G. A four-column theory for the origin of the genetic code: tracing the evolutionary pathways that gave rise to an optimized code. *Biol Direct*. **2009**, *24*, 4-16.
24. Van der Gulik, P.; Massar, S.; Gilis, D.; Buhrman, H.; Rooman, M. The First Peptides: The evolutionary transition between prebiotic amino acids and early proteins. *J. Theor. Biol.* **2009** *261*, 531-539.
25. Yanagawa, H.; Ogawa, Y.; Kojima, K.; Ito, M. Construction of protocellular structures under simulated primitive earth conditions. *Orig. Life Evol. Biosph*. **1988**, *18*, 179-207.
26. Ikehara, K.; Omori, Y.; Arai, R.; Hirose, A. A novel theory on the origin of the genetic code: a GNC-SNS hypothesis. *J. Mol. Evol*. **2002**, *54*, 530-538.