

Review

Not peer-reviewed version

How was the “Chicken and Egg Relationship” between Gene and Protein formed?

[Kenji Ikehara](#) *

Posted Date: 28 June 2023

doi: 10.20944/preprints202306.1878.v1

Keywords: chicken and egg relationship; origin of gene; origin of protein; GADV hypothesis; origin of life



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Review

How was the “Chicken and Egg Relationship” between Gene and Protein formed?

Kenji Ikehara ^{1,2}

¹ G&L Kyosei Institute, In the Keihanna Academy of Science and Culture (KASC), Keihanna interaction plaza, Lab. Wing 3F, 1-7 Hikaridai, Seika-cho, Souraku, Kyoto 619-0237, Japan; ikehara@cc.nara-wu.ac.jp; Tel.: +81-774-73-4478

² International Institute for Advanced Studies, 9-3 Kizugawadai, Kizugawa, Kyoto 619-0225, Japan

Abstract: It has not been reasonably explained still now, how the “chicken and egg relationship” between gene and protein was formed, although RNA world hypothesis was proposed to aim at explaining formation process of the relationship about 40 years ago. On the other hand, GADV hypothesis advocates that life emerged after five members of the fundamental life system; cell structure, metabolism, tRNA, genetic code and gene, were piled up one by one onto immature [GADV]-proteins, which were produced by random joining of [GADV]-amino acids on the primitive Earth. Furthermore, formation process of the first (GNC)_n RNA gene encoding a mature [GADV]-protein can be also reasonably explained according to the hypothesis as follows. Double-stranded (ds)-(GNC)_n RNA could be produced by complementary strand synthesis of a single-stranded (ss)-(GNC)_n RNA, which was formed by random joining of GNC anticodons carried by four types of AntiC-SL tRNAs. Successively, the first (GNC)_n RNA gene was generated upon maturation of an immature [GADV]-protein, which was produced through expression of codon sequence on one strand of ds-(GNC)_n RNA, as accumulating necessary base substitutions and raising a weak catalytic activity of the immature protein to a more active state. It is considered that, consequently, a gene encoding a mature protein, which has the “chicken and egg relationship”, was generated for the first time. Therefore, it can be concluded that the “chicken and egg relationship” was built not by formation of either one, gene or protein, but by maturation of an immature gene, which was lead by improvement of catalytic activity of an immature protein. That is, it can be concluded that the reason, why it was unable to give an answer to the question on the formation process of the “chicken-egg relationship” thus far, is because it was an unanswered question.

Keywords: chicken and egg relationship; origin of gene; origin of protein; GADV hypothesis; origin of life

1. Introduction

It is quite difficult to answer to the question “Which came first? The chicken or the egg?”. The reason is because a hen is indispensable for an egg to be laid down and on the contrary an egg is indispensable for a hen to be born. Therefore, it is often stated that the “chicken-egg relationship” is one representative example, which it is quite difficult or even it might be impossible to solve because of a kind of circular argument.

1.1. “Chicken and egg relationship” between gene and protein

Gene and protein have the “chicken and egg relationship” as similarly to the genuine “chicken and egg” (Figure 1) [1-4]. Needless to state, gene and protein simply described in this article are visible things, which are saved into various gene and protein databases. In addition, as well known, both gene and protein cannot be synthesized by random joining of nucleotides and amino acids, respectively [5]. In other words, gene and protein mean a mature gene and a mature protein, respectively. Every mature protein is always synthesized through expression of the corresponding gene. Therefore, any protein cannot be produced in the absence of gene. On the other hand, at least some proteins are always required to express genetic information. Therefore, genetic information carried by a gene cannot be expressed in the absence of protein. How was the “chicken-egg

relationship” between gene and protein formed? It is the purpose of this article to give an answer to the question.

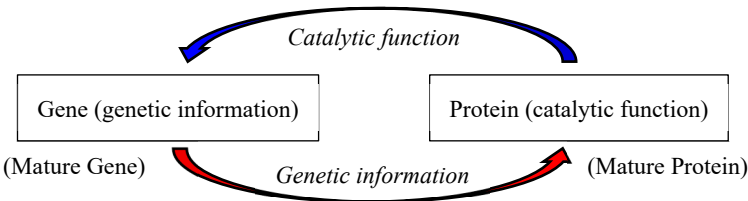


Figure 1. The “chicken-egg relationship” is observed between gene and protein, because expression of genetic information requires catalytic function of protein (blue arrow). On the contrary, genetic information is necessary to produce protein with catalytic function (red arrow).

1.2. Definitions of mature gene/protein and immature gene/protein

Gene described simply in the above sentences naturally means a gene encoding a mature protein. However, it needs to suppose immature genes encoding an immature protein to make clear the formation process of mature genes encoding a mature protein and to give a correct answer to the question, “which came first, gene or protein?”. Furthermore, there are large differences among some types of immature proteins as described below. Therefore, in addition to abbreviations of several words used in this article (Table 1), definitions of genes and proteins are summarized in Table 2 for readers to make it easier to understand the content of this article and to address the key points about mature genes, immature genes and so on.

Mature genes are genes encoding a mature protein, which is formed upon maturation of an immature gene. Mature proteins are proteins, which are encoded by a gene and are frequently called as a precision polymer machine. Note that the maturation of an immature gene became possible, which was always lead by enhancement of a weak activity on surface of an immature protein [5]. Only base sequences of mature genes and amino acid sequences of mature proteins are usually saved in various databases.

There are two types of immature genes [5]. Immature gene 1, is the first double-stranded (ds)-RNA, which was formed by random joining of GNC anticodons carried by anticodon stem loop (AntiC-SL) tRNA. Therefore, an immature gene 1 encodes an immature [GADV]-protein having a random [GADV]-amino acid sequence (Table 2).

In the second type of immature gene or immature gene 2, there are three types of GC-NSF(a)s; GC-NSF(a) 2-1: GC-rich (GNC)_n-NSF(a) encoding an amino acid sequence consisted of four [GADV]-amino acids in GNC code era. GC-NSF(a) 2-2: GC-rich (SNS)_n-NSF(a) encoding amino acid sequence composed of 10 types of amino acids in SNS code era. GC-NSF(a) 2-3: Nonstop frame on antisense strand of GC-rich (NNN)_n genes (GC-NSF(a)), which was first recognized as a field for formation of an entirely new gene in universal genetic code era [6] (Table 2).

There are three types of immature proteins [5]. Immature protein 0: an immature [GADV]-protein formed as peptide aggregates, which were produced by direct random joining of [GADV]-amino acids or with activated [GADV]-amino acids. Immature protein 1: an immature [GADV]-protein, which was produced by expression of one strand of the first ds-(GNC)_n RNA. Immature protein 2: an immature protein, which was produced by expression of three types of GC-NSF(a)s or GC-NSF(a) 2-1, GC-NSF(a) 2-2 and GC-NSF(a) 2-3 (Table 2).

Table 1. Abbreviations of several terms appeared in this article.

Abbreviation	Meaning	Abbreviation	Meaning
GADV	G: Gly, A: Ala, D: Asp, V: Val	NSF(a)	Nonstop frame on antisense strand
GNC	N: either one of four bases (U, C, A, G)	AntiC-SL	Anticodon stem-loop
SNS	S: either one of G and C		

Table 2. Definitions of mature gene, mature protein, immature gene and immature protein are given.

Mature gene and mature protein	
Mature gene	A gene encoding a primary structure of a mature protein with an elaborate structure
Mature protein	A protein with a tertiary structure, which is encoded by a mature gene
Immature gene and immature protein	
Immature gene: A gene encoding a protein 0 th -order structure or an amino acid sequence of an immature protein	
Immature gene 1	The first double-stranded (ds)-RNA, on which a protein 0 th -order structure is written
Immature gene 2	A nonstop frame on antisense strand of three types of GC-rich genes
GC-NSF(a) 2-1	GC-rich (GNC) _n -NSF(a), in which a protein 0 th -order structure is written, in GNC code era
GC-NSF(a) 2-2	GC-rich (SNS) _n -NSF(a), in which a protein 0 th -order structure is written, in SNS code era
GC-NSF(a) 2-3	GC-NSF(a), in which a protein 0 th -order structure is written, in universal genetic code era
Immature protein: An immature protein produced under a protein 0 th -order structure	
Immature protein 0	A protein as peptide aggregate produced by random joining of [GADV]-amino acids
Immature protein 1	An immature [GADV]-protein produced by expression of one strand of ds-(GNC) _n RNA
Immature protein 2	An immature protein produced by expression of three types of GC-NSF(a)s, 2-1, 2-2, 2-3

Naturally, immature protein 0, which was produced in the absence of ds-RNA, could not be matured to a mature [GADV]-protein. On the contrary both an immature protein 1 and 2, which were encoded by either one of strands of ds-RNA and by a GC-NSF(a) of ds-RNA/DNA gene, could evolve to a mature protein, respectively. Note that all the three types of immature proteins, 0, 1 and 2, were produced under respective protein 0th-order structures [5,7].

1.3. The “chicken and egg relationship” between gene and protein in RNA world hypothesis

As a matter of course, elucidation of formation process of the genetic system composed of gene and protein holds one important key for solving the mystery of the origin of life. However, the “chicken-egg relationship” has become one large obstacle against solving the mystery. At just the time, RNAs, which have a catalytic activity similarly to an enzyme and a chemical structure similar to DNA, or ribozymes, were discovered [8,9]. The discoveries triggered proposition of the RNA world hypothesis on the origin of life [1]. Therefore, it was considered that the mystery of the origin of life might be solved, if RNA itself can synthesize RNA or RNA can be self-replicated. In fact, many researches studying in the field of the origin of life have studied at the basis of the RNA world hypothesis. However, before long it has become to understand that there are many fatal flaws in the RNA world hypothesis [5,10-12]. For example,

1. It is quite difficult to synthesize nucleotides and RNA with prebiotic means.
2. It is also difficult or probably impossible for RNA to self-replicate, although some RNAs can synthesize oligonucleotides [13,14].
3. Genetic information for protein synthesis never be written into RNA strand in the absence of protein.

In fact, the mystery of the origin of life has not been solved thus far due to the above reasons.

These mean that the origin-life investigation must be started over again from scratch, if the mystery of the origin of life cannot be solved by the RNA world hypothesis. In parallel, it was found that the questions, as “how was the genetic system composed of gene and protein formed under random processes on the primitive Earth?” and “how did life emerge on the primitive Earth?”, must be reconsidered from the beginning.

1.4. *The reasons why the “chicken and egg relationship” between gene and protein could not be solved thus far*

Consider here what makes it difficult to solve the “chicken-egg relationship” between gene and protein and what needs to make clear the formation process of the relationship.

1. The “chicken-egg relationship” is observed between every mature gene and every corresponding mature protein, which is frequently called as a precision polymer machine.
2. On the other hand, both such a mature gene and a mature protein cannot be produced by random joining of the respective monomers as nucleotides and amino acids [5,15].

Those would be the reasons why it was unable to give an answer to the question, which came first, gene or protein?.

1.5. *The “chicken-egg relationship” between gene and protein was established upon formation of a mature gene*

The formation process of the “chicken-egg relationship” should be made clear, if the formation process of mature genes, which encode a mature protein with an elaborate structure, could be understood, because the formation process of the “chicken-egg relationship” must be directly related to the acquisition of amino acid sequence information of a mature protein by the respective genes. Thus, we have noticed that the formation process of the “chicken-egg relationship” between genes and proteins could be reasonably explained through formation process of mature genes.

It is necessary to understand the following three types of formation processes of new genes in order to give an answer to the question about the “chicken and egg relationship”. Formation processes of mature genes are classified into the following three.

1. How was the first gene generated?
2. How were entirely new genes except the first gene, which do not possess any meaningful homology with any other previously existed genes, formed?
3. How were new genes having meaningful homology with any other previously existed genes formed?

Then, the formation processes of the three types of new genes are explained in order in the following Sections. Note that entirely new genes including the first gene were formed by maturation of immature gene through either one way for the above two formation processes 1 and 2 of a mature gene. On the contrary, new homologous genes encoding a mature protein were formed by transition of an original gene without using an immature gene. The “chicken-egg relationship” between gene and protein was established at the moment of completion of the respective mature genes.

2. How did the first gene acquire an amino acid sequence information of a protein?

2.1. *The key points for the formation of the first (GNC)_n gene*

The formation process of the first gene is explained using schematic drawing of the whole process from chemical evolution (step 1) to the emergence of life (step 8) via formation of the first gene (step 7) (Figure 2) [5]. The first gene encoding a mature protein was generated through the process from chemical evolution, [GADV]-microsphere formation, formation of a primeval metabolic system, formation of AntiC-SL tRNA or prototype tRNA and establishment of GNC code. The process proceeded owing to various functions expressed on immature [GADV]-proteins, which were produced by random joining of [GADV]-amino acids.

What are key points for the formation of the first (GNC)_n gene shown in Figure 2? (1) Water-soluble globular immature [GADV]-protein 0 could be produced even by random joining of [GADV]-amino acids owing to the protein 0th-order structure or [GADV]-amino acid composition, which satisfies the four conditions (hydrophobicity/hydrophilicity, α -helix, β -sheet and turn/coil formabilities) for water-soluble globular protein formation [16-18]. (2) Immature protein 1 was synthesized through transcription of ds-RNA and translation of the transcript. Note that immature [GADV]-proteins produced by expression of random (GNC)_n codon sequences are substantially the same with immature [GADV]-proteins produced by direct random joining of [GADV]-amino acids.

Therefore, both immature [GADV]-proteins 0 and 1 could be folded into the respective globular structures in water at a high probability [16-18]. In other words, the first gene was generated through the process of maturation of an immature protein synthesized under a primeval transcription/translation system including many catalytic functions of such immature [GADV]-proteins, which were produced in the absence of any genetic function [5].

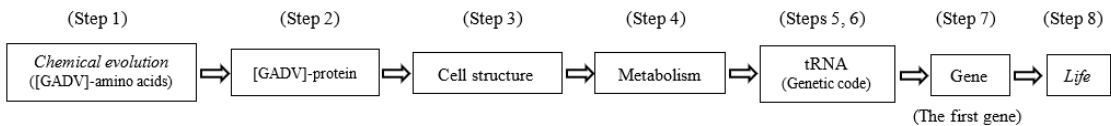


Figure 2. Possible evolutionary steps from chemical evolution (step 1) to the emergence of life (step 8) via formation of the first gene (step 7), which is deduced from GADV hypothesis, is shown in a linear manner.

Although both immature protein 0 and immature protein 1 could be folded into immature and flexible but water-soluble globular structure [5], there is a crucial difference between immature protein 0 and immature protein 1, because amino acids are only randomly arranged from protein 0th-order structure in protein 0 and on the contrary an immature protein 1 was produced based on the ds-RNA (Table 2). Therefore, there was no possibility for protein 0 to evolve to a mature protein, although it was possible for protein 1 to evolve to a mature protein by using memorizing ability of base changes of ds-RNA. Thus, life could emerge about 4 billion years ago owing to indirect function of protein 0 and direct function of protein 1 (Figure 2).

2.2. Evolutionary process of synthesis of immature protein 0

Next, explain some processes to formation of the first gene in more detail. Only synthesis of an immature protein 0 could be carried out by direct random joining of [GADV]-amino acids, and by random joining of [GADV]-amino acids with activated amino acids, such as [GADV]-AMPs, [GADV]-3'ACC5', [GADV]-amino acids carried by nonspecific 3'ACC5'-AntiC-SL RNAs. Such uses of activated [GADV]-amino acids were the results that [GADV]-microspheres searched for mechanisms producing immature [GADV]-proteins with a higher [GADV]-amino acid content and a higher catalytic activity than before [5].

2.3. From formation of single-stranded (GNC)_n RNA to generation of the first (GNC)_n gene

Successively, essentially random polymerization of [GADV]-amino acids were also carried out to produce an immature [GADV]-protein 1. That is, the immature [GADV]-protein 1 was synthesized using ds-(GNC)_n RNA, which was formed by complementary strand synthesis of ss-(GNC)_n RNA, under a prototype of transcription-translation system with immature [GADV]-proteins and nonspecific AntiC-SL tRNAs (Figure 3) [5].

Evolution of a weak catalytic activity, which appeared on a surface of the immature [GADV]-protein 1, to a mature [GADV]-protein with a higher catalytic activity generated the first mature gene (Figure 3). Thus, the first gene encoding one mature protein was generated owing to many immature [GADV]-proteins 0, which were produced by random joining of [GADV]-amino acids, independently of mature genes.

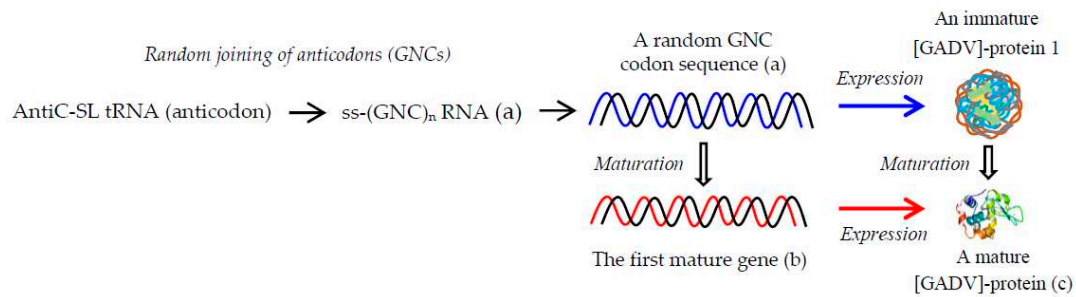


Figure 3. After single-stranded (ss)-(GNC)_n RNA (a: protein 0th-order structure) was synthesized by random joining of anticodons, GNCs, which were carried by AntiC-SL RNAs [5], double-stranded (ds)-random (GNC)_n RNA (a) was produced through complementary strand synthesis of the ss-(GNC)_n RNA (a). Finally, the first ds-(GNC)_n RNA gene (b: protein primary structure) encoding an amino acid sequence of a mature protein (c: protein tertiary structure) was generated by maturation of a (GNC)_n sequence (a) on one strand of the ds-(GNC)_n RNA as accumulating necessary base substitutions. (a), (b) and (c) indicate that they are related to protein 0th-order, primary and tertiary structures, respectively.

Of course, the “chicken and egg relationship” can be observed not only between the first gene and the protein, which was produced by expression of the first gene, but also in all pairs between a mature gene and the corresponding mature protein, such as even between modern genes and modern proteins. So, next, it is explained how entirely new genes and homologous genes were formed, so that the reason, why the “chicken and egg relationship” is observed between all pairs of mature genes and mature proteins, which were produced by genetic expression of the corresponding genes, can be understood.

2.4. Formation process of entirely new genes after generation of the first gene

Three types of GC-NSF(a)s 2-1, 2-2, 2-3 were used to generate entirely new genes in every genetic code era (Table 2). (1) ds-(GNC)_n genes in GNC primeval genetic code era. (2) ds-(SNS)_n genes in SNS primitive genetic code era. (3) GC-rich genes having a codon sequence similar to (SNS)_n sequence in the universal (standard) genetic code era [5,16-18]. The immature proteins 2, which were produced by expression of antisense sequences of the respective three types of GC-rich genes, could be folded into water-soluble globular structures at a high probability, because the immature proteins could satisfy the four or six conditions for water-soluble globular protein formation [5,16-18]. Therefore, mature proteins, which were formed through maturation of the respective immature proteins, are entirely new [GADV]-proteins in GNC code era, entirely new SNS-coding proteins in SNS code era and entirely new proteins encoded by GC-rich genes in universal genetic code era, respectively (Figure 4) [5].

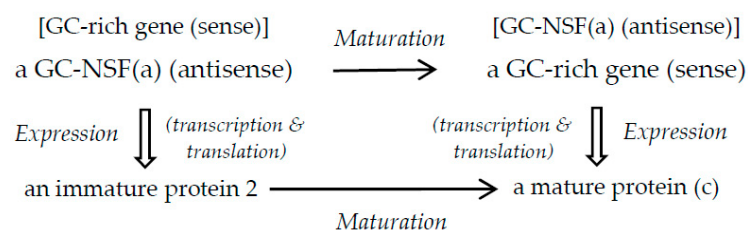


Figure 4. Generation process of an entirely new gene after formation of the first ds-(GNC)_n RNA gene. Entirely new genes were generated by maturation of a nonstop frame on antisense strand (NSF(a)) of three types of GC-rich genes, (GNC)_n genes in GNC code era, (SNS)_n genes in SNS code era and ordinary GC-rich genes in the universal (standard) genetic code era. After maturation, an original GC-NSF(a) and an original GC-rich gene on sense strand became an entirely new GC-rich gene and a new GC-NSF(a), respectively. (a), (b) and (c) indicate that they are related to protein 0th-order, primary and tertiary structures, respectively.

Thus, many entirely new genes were generated through maturation of an immature protein produced by expression of immature genetic information, which was written in a nonstop frame on antisense strand of three-types of GC-rich gene (pan-GC-NSF(a)) (Figure 4) [5]. Naturally, all the maturation process are essentially the same as the case of formation of the first gene except usage of ds-RNS or of GC-NSF(a) (Figure 3). This means that the “chicken and egg relationship” between one gene and the corresponding protein was always established at the moment when a gene encoding a mature protein was completed by maturation of a ds-RNA and a GC-NSF(a) encoding an immature protein, respectively (Figure 4). All entirely new genes were formed owing to immature genes encoding an entirely new immature protein, although it is natural when the first gene was generated in the absence of gene [5]. Thus, immature proteins played the lead role in generating two types of entirely new genes, one is the first gene and the others are all entirely new genes formed after generation of the first gene (Figures 3 and 4). Then, the propensity to evolve gradually from an immature protein with a weak catalytic activity to a mature protein with a high activity as accumulating necessary mutations was the motive force, which lead to formation of all types of entirely new genes [5].

2.5. Formation process of producing homologous genes

It should be noted that the process generating a homologous gene as accumulating necessary mutations on sense codon sequence after gene duplication [19] is a transition process from an original gene encoding a mature protein to a new gene encoding another mature protein homologous to the original protein (Figure 5). That is, the process is a simple transition from a mature protein to another new mature protein, although intermediates, which appeared during the transition process, might pass through a kind of immature state with some flexibility to adjust the original catalytic site to a new substrate.

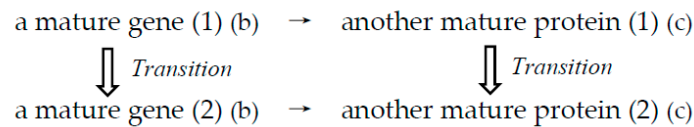


Figure 5. Formation process of a gene homologous with an original gene. This is a transition process from an original mature gene (1) to a new homologous gene (2). Both the original mature gene (1) encoding a mature protein (1) and another new gene (2) encoding another mature protein (2) with some homology to the original protein have the “chicken-egg relationship” and are visible in various databases of genes and proteins. (b) and (c) indicate that they are related to primary and tertiary structures, respectively.

3. Discussion

It is described in this article that formation process of the “chicken-egg relationship” between gene and protein can be reasonably explained from the standpoint of the GADV hypothesis. That is, according to the GADV hypothesis, it is considered that the first gene encoding a mature [GADV]-protein could be generated through a maturation process from an immature [GADV]-protein 1, which was produced by expression of one strand of ds-(GNC)_n RNA (Figure 3). This means that the question, “which did first arise, gene or protein?” was an unanswered question. The formation process of the “chicken-egg relationship” could not be understood forever, as long as the question is forthrightly considered.

3.1. The “chicken-egg relationship” of gene /protein was formed by maturation process of an immature protein

The first “chicken-egg relationship” was established at just the moment, when the first gene encoding a mature protein was formed owing to an immature protein 1 (Figure 3). However, the

“chicken-egg relationship” cannot be observed between the immature gene and the corresponding immature protein (Figure 6).

All other new genes encoding a mature protein with a high catalytic activity were formed from three types of GC-NSF(a)s 2 ((GNC)_n(a) 2-1, (SNS)_n(a) 2-2 and ordinary GC-NSF(a) 2-3) encoding an immature protein 2, through maturation of an immature gene (Figures 3 and 4), or by transition from a parental mature gene to a daughter mature gene (Figure 5). Therefore, the “chicken-egg relationship” always arose at the time when a gene encoding a mature protein was formed. In this way, various genes encoding a mature protein were generated and life emerged when genes necessary for life to live could be equipped. The wonderful present Earth, on which versatile organisms are inhabiting, has been formed as those organisms have been accumulating many versatile genes under mechanisms generating various mature genes. Consequently, many mature genes and many mature proteins, which have the “chicken-egg relationship”, have been formed and only data of such genes and proteins are saved in the present gene/protein databases. Inversely stating this, intermediate genes/proteins in a maturation process could not be generally seen (Figure 6), although it might become possible to see even those intermediate genes/proteins by future investigation of modern databases of genes and proteins. The invisible process starting from an immature gene and proceeding to a mature gene have made it difficult to give an answer to the question or formation process of the “chicken-egg relationship” thus far (Figure 6). The reason, why it has been considered that it is quite difficult or might be impossible to explain the formation process of the “chicken-egg relationship” between gene and protein, is because the maturation process from an immature gene to a mature gene have not been noticed thus far. GADV hypothesis has succeeded to visualize the invisible process and to explain the formation process of the “chicken-egg relationship”. This also would support the validity of the GADV hypothesis [5,16,17].

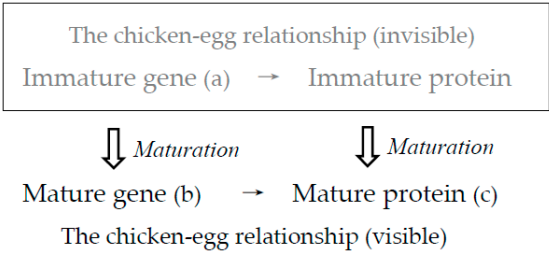


Figure 6. As shown in Figure 5, such entirely new genes were generated by maturation from an immature gene (a: protein 0th-order structure) to a mature gene (b: protein primary structure) encoding a mature protein (c: protein tertiary structure). However, both an immature gene and an immature protein described in gray letters are usually invisible until a mature gene encoding a mature protein has been completed. Inversely stating this, only mature genes and mature proteins, which have so-called the “chicken-egg relationship”, can be seen in gene-protein databases. (a), (b) and (c) indicate that they are related to protein 0th-order, primary and tertiary structures, respectively.

3.2. The “chicken-egg relationship” observed among 6 members consisting of the fundamental life system

Actually, the “chicken-egg relationship” can be observed even in all combinations among the 6 members used in extant organisms (Table 3). The reason, why the “chicken-egg relationship” can be observed in all the combinations, is explained below, although it is shown here only the “chicken-egg relationships” between gene and the other four members except protein and between protein and the other four members except gene to simplify the discussions.

Table 3. Combinations, among which the “chicken-egg relation ship” can be seen among 6 members consisting of the fundamental life system used by extant organisms. Circles indicate the combinations, between which the “chicken-egg relation ship” can be seen.

	Gene	Genetic code	tRNA	Metabolism	Cell structure	Protein
Gene	*					
Genetic code	○	*				
tRNA	○	○	*			
Metabolism	○	○	○	*		
Cell structure	○	○	○	○	*	
Protein	○	○	○	○	○	*

3.3. The “chicken-egg relationships” between gene and other four members except protein

1. Gene and genetic code: Genetic code is meaningless and useless in the absence of gene. Genetic information cannot be written in RNA or DNA strand in the absence of genetic code. This means that gene never be formed without genetic code.
2. Gene and tRNA: Genetic information in a gene cannot be expressed in the absence of tRNA. On the other hand, tRNA for translation of genetic information is useless in the absence of gene.
3. Gene and metabolism: Genetic information carrier, RNA or DNA, cannot be produced in the absence of metabolism, in which nucleotides are synthesized. Metabolic system cannot be formed in the absence of gene, because bio-catalysts or mature enzymes cannot be synthesized in the absence of gene.
4. Gene and cell structure: Cell structure cannot be constructed in the absence of gene, because membrane proteins cannot be produced without gene. Genes would be dispersed if cell membrane is absent, even if genes could be formed.

3.4. The “chicken-egg relationships” between protein and other four members except gene

1. Protein and genetic code: Protein cannot be synthesized without genetic code, because genetic information cannot be translated without genetic code. Genetic code mediating between genetic information and protein is useless in the world in the absence protein.
2. Protein and tRNA: tRNA cannot be produced without protein (enzyme). Protein cannot be synthesized without tRNA, because genetic information in gene cannot be translated into amino acid sequence in the absence of tRNA.
3. Protein and metabolism: Metabolic system cannot be driven without protein (enzyme). Protein cannot be produced in the absence of metabolism, because amino acids, of which protein is composed, are not supplied through metabolic system.
4. Protein and cell structure: Cell structure cannot be constructed in the absence of protein, because membrane using proteins cannot be produced in the absence of proteins. Proteins would be dispersed without cell structure, even if proteins could be synthesized.

It must be emphasized here that the “chicken-egg relationships” observed among all combinations of 6 members can be explained based on the “chicken-egg relationships” observed between mature genes and mature proteins.

References

1. Gilbert, W. The RNA world, *Nature* **1986**, 319, 618.
2. Robertson, M.P.; Joyce, G.F. The origins of the RNA world. *Cold Spring Harb. Perspect. Biol.* **2012**, 4:a003608.
3. Sankaran, N. How the discovery of ribozymes cast RNA in the roles of both chicken and egg in origin-of-life theories. *Stud. Hist. Philos. Biol. Biomed. Sci.* **2012**, 43, 741-750.
4. Sankaran, N. The RNA World at Thirty: A Look Back with its Author. *J. Mol. Evol.* **2016**, 83, 169-175.

5. Ikehara, K. *Towards Revealing the Origin of life.—Presenting the GADV Hypothesis*; Springer Nature, Gewerbestrasse: Cham Switzerland, **2021**.
6. Ikehara K, Amada F, Yoshida S.; Mikata, Y.; Tanaka, A. A possible origin of newly-born bacterial genes: Significance of GC-rich nonstop frame on antisense strand. *Nucl. Acids Res.* **1996**, 24:4249–4255.
7. 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.
8. 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.
9. 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.
10. Shapiro, R. A replicator was not involved in the origin of life. *IUBMB Life* **2000**, 49, 173–176.
11. Luisi, P.L. An open question on the origin of life: the first forms of metabolism. *Chem. Biodivers.* **2012**, 9:2635–2647.
12. Ikehara, K. Life Emerged from [GADV]-Protein World, but Not from RNA World!? *Preprints* doi:10.20944/preprints201712.0170.v1
13. Joyce, G.F.; Szostak, J.W. Protocells and RNA Self-Replication. *Cold Spring Harb. Perspect. Biol.* **2018**, 10:a034801.
14. Salditt, A.; Karr, L.; Salibi, E.; Le Vay, K.; Braun, D.; Mutschler, H. Ribozyme-mediated RNA synthesis and replication in a model Hadean microenvironment. *Nat Commun.* **2023**, 14:1495.
15. Dill, K.A. Dominant forces in protein folding. *Biochemistry* **1990**, 29:7133–7155.
16. 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.
17. 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. *J. Biosci*
18. Ikehara, K. Possible steps to the emergence of life: The [GADV]-protein world hypothesis. *Chem. Rec.*, **2005**, 5:107–118.
19. Ohno S. *Evolution by gene duplication*, Springer, Heiderberg, Germany, **1970**.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.