Open Peer Review on Qeios

Numeric Structure of Genetic Code in Natural Evolution: Energy Grounds

Andrei Moldavanov

Funding: The author(s) received no specific funding for this work.

Potential competing interests: The author(s) declared that no potential competing interests exist.

Abstract

It is shown that on the one hand, the track of energy evolution for open systems with conservation links possesses a few pairs of outstanding bifurcation points only, which forms the finite phase structure of energy evolution. On the other hand, for cells, the same process of energy evolution can be described by a limited number of major interchangeable mechanisms of energy metabolism. On this basis, we develop a novel approach for the origin of a numerical skeleton for genetic code (*GC*). It is revealed that, in course of energy evolution, an original*6*-phase structure can be reduced to a *4*-phase one. Alternatively, we observe that energy evolution for cells can be also presented as the random coexistence of three interchangeable metabolism pathways - glycolysis, photosynthesis, and oxidative metabolism. It assumes the existence of the dual essence of the same process of energy evolution. Then, a mathematically correct way to describe energy evolution is to combine two different characteristics of evolution through a combinatorial approach. From a combinatorics standpoint, we have a classic problem for variation of *4* "balls" over *3* "bins| which yields *64* possible scenarios for codon arrangement. Similarly, the number of necessary amino acids at ignorance of the order parameter is *20*. As a result, the suggested approach intuitively combines a few chief known principles of natural evolution into one consistent way – (a) mathematical foundation for the spectral structure of system energy evolution; (b) existence of different but compatible mechanisms of energy metabolism in cells; (c) forming of the numerical structure of *GC* with the magic numbers *3, 4, 20,* and *64*.

Keywords: *natural evolution, evolutionary unit, genetic code, codon, triplet.*

1. Introduction

1.1. Generals

Though it is **well known** that the processes of natural evolution possess some identified and construed common features[1], a commonly accepted physical and mathematical theory for biological development still is not built[2].

Meantime, the scientific community generally agrees on the existence of the observable consecutive steps in the transformation of non-animated things to real life. The two biggest and most accepted steps are abiogenesis and biogenesis. Chronologically, the first step (abiogenesis) basically deals with the chemistry transformations. The second one (biogenesis) incorporates the time interval when the Darwin-Wallace mechanism of life's expansion has become the main driving force for changes[3]. The presence of other undoubtedly visible steps is under regular scientific review[4]. In some connection, the hypothetical existence of the more or less demarcated points in the evolution process might be related to the well-known practical example of evolutionary stepping which is a phenomenon of a genetic code (*GC*) that surprisingly works in the continuous struggle with the intrinsic noise while accurately and efficiently translates discrete inheritance information[5].

It is well-agreed that an essence of *GC* is to instruct the cell on how to build *DNA*. Then, it is logical to believe that *GC* was created and naturally integrated into the primordial environment prior to *DNA* arrival. In this sense, the structure of *GC* has a direct connection to the physical laws that ruled on the Earth since the beginning[6]. The opposite viewpoints, mainly, the associate appearance of *GC* with the coming of the building material for its working, *i.e.,* the amino acids[7], for example, the theory of a "frozen accident"[8]. In any case, there are serious arguments to state that the mechanism for forming the original *GC* has started as early as in abiogenesis and becomes fully operational in biogenesis[6].

These days, there is a consensus within the scientific community that in its structure, the alphabet of *GC* consists of *4* letters (nucleotides) adenine, cytosine, guanine, and thymine[9]. These letters are randomly united in the not overlapping[10] triplets (codons) with a total number of *64* possible codon combinations forming the vocabulary of *GC* which manages the assembly of proteins from *20* necessary amino acids[11]. Currently, there are three main theories dealing with the origin and evolution of *GC*. These theories are (1) the error minimization theory considering natural adjustment as a tool for evaluation of impact in mutation and translation; (2) the stereochemical theory based on physicochemical affinity between amino acids and anticodons; and (3) the coevolution theory which states that the code structure coevolved together with biosynthesis of amino acid[7]. Despite substantial contribution, the above theories do not provide sufficient arguments in favour of the existing numerical basis of *GC* as *3, 4, 20, 64*[12].

So, the major purpose of this report is to review the energy (physical) and mathematical grounds which by working together could favour the known numerical structure of *GC*. We find that such grounds can create an energy carcass for further biochemical layers which enrich the energy foundation with a variety of features inherent to actual life.

Should be noted that doing this presentation, we will refrain from any comments on the biochemical factors which may or may not convey the discovered energy transformations.

1.2. Energy evolution of open system

From the said above, it looks reasonable to consider *GC* as an integral part of the general long-term evolution of nature, then the sought driving force of *GC* evolution could be based on the as ubiquitous and permanent elements of Earth's environment through the entire natural history as possible.

One of the most general environmental relations we typically deal with is an energy conservation law. Its differential form is also known as an energy continuity equation (*ECE*) that in the non-relativistic approximation can be written as

$$\frac{\partial \varepsilon}{\partial t} + \nabla \cdot \mathbf{J} = 0 \qquad (1.a)$$

where J is the flux of energy, ε is the energy volume density, t is time, and ∇ is the Nabla operator.

Then, a suitable mathematical model (1.b) for energy evolution of an open entity is an infinite system of energy links (1.a) in the assumption of permanent energy exchange with the thermal bath of unlimited capacity[13].

Note, that the discussed approach is based on an abstraction of the conserved energy link (*CEL*), which incorporates all energy exchange scenarios including those in which change in the system state is projected onto an energy axis no matter what sort of exchange process or combination of processes it deals with.

So, based on the above, system (1.b) is

$$\begin{cases} \frac{\partial \varepsilon_1}{\partial t} + \nabla \cdot \mathbf{J}_1 = 0 \\ \frac{\partial \varepsilon_2}{\partial t} + \nabla \cdot \mathbf{J}_2 = 0 \\ \quad \ldots \\ \frac{\partial \varepsilon_n}{\partial t} + \nabla \cdot \mathbf{J}_n = 0 \\ \quad \ldots \end{cases} \qquad (1.b)$$

with solution

$$\Upsilon(y) = y - y\ln y \qquad (2)$$

where Υ is the average efficiency of total bidirectional energy exchange between the system and the environment shown in Fig. 1, $y = J/J_0$ is *the* rate of energy exchange*,* normalization constant $J_0 > 0$, $n \to \infty$ [13].

As $\Upsilon(y)$ is bounded in the range[0, e], it features the discrete spectrum

$$y_n = exp[\pm \frac{1}{n}] \qquad (3)$$

where e is the base of the natural logarithm.

Further, a uniform probability distribution (*UPD*) will be used for all involved random quantities.

More detail of the presented model can be found in[4][13][14][15][16].

2. Physical grounds of *GC* origin

2.1. Nonuniform bifurcation in spectral nodes

The same result (3) can be found using the apparatus of theoretical mechanics. Then, the discrete nodes (3) can be additionally considered as the points of bifurcation (*PoB*) for the process of energy exchange (1.b) that separate qualitatively different areas of the solution.

Further analysis confirmed that the nodes with $n \le 3$ clearly demonstrate bifurcation behaviour [4][14]. However, the nodes with $n > 3$ do not reveal similar properties to a sufficient extent. At the very least, if the bifurcation behavior for $n > 3$ does exist, it looks degenerate compared with the nodes at $n \le 3$ [4]. Therefore, we have to acknowledge the different roles of nodes (3).

So, below we will be ignoring individual contributions to the bifurcation performance for the nodes with $n > 3$, however, keep chance for their collective contribution, which, in this case, will be limited by the two middle phases of energy evolution only (see *P3* and *P4* in Fig. 1).

Consequently, the structure of energy evolution is essentially defined by the *3* pairs (in respect of $y = 1$) of *PoB* (3) only that together with the stationary point (*SP*) at $y = 1$ compose appropriate interpoint phases *Pi* as the areas of unlikely energy conditions, here $i = 1, 2 \ldots 6$ (Fig. 1).

From [4], based on the used assumption of *UPD*, the design ratio

$$\vartheta(n) = \frac{p_{ev}}{p_{nev}} = \frac{|ln\, y|}{1 - |ln\, y|} = \frac{1}{n-1},\ n \neq 1 \qquad (4)$$

where p_{ev} deals with the probability of evolutionary scenario while p_{nev} with the probability of non-evolutionary scenario (in terms of [17]).

To compare physical conditions in the discovered phases of energy evolution *Pi*, look at Table 1 below. This table was created based on (2,4) and a simple calculation of the sign for $v = d\Upsilon/dy$.

Physical essence of phases

Table 1. Unique physical conditions in the phases of energy evolution.

Phase #	v	ϑ	ζ
Phase 1	*Positive*	$1 < \vartheta < \infty$	$p_{ev} > p_{nev}$
Phase 2		$0.5 < \vartheta < 1$	$p_{nev} > p_{ev}$
Phase 3		$0 < \vartheta < 0.5$	$p_{ev} = 0$
Phase 4	*Negative*		
Phase 5		$0.5 < \vartheta < 1$	$p_{nev} > p_{ev}$
Phase 6		$1 < \vartheta < \infty$	$p_{ev} > p_{nev}$

Sign of v, as well as the value of ratios ϑ and ζ, is shown.

We see that the combination of physical conditions determined by the factors v, ϑ, ζ is unique in each phase *Pi*.

It is worth noting here that the trailing phases *P4, P5, and P6* fit the stage of reverse energy development when $d\Upsilon/dy < 0$ as well as circumnavigation of instant energy efficiency $\delta\Upsilon$ changes to the opposite[15].

Note that within the presented theory, exchange energy E_n at arbitrary spectral node n is the product $y_n \cdot \Upsilon_n$. It follows from the physical meaning of y and Υ as highlighted in the introduction. Then, in the *k*-базисе, there is a fundamental ratio between the first three energy harmonics

$$E_1 = \frac{1}{2} E_2 E_3 \qquad (5.a)$$

which also can be given as geometric mean

$$E_1 = \left(\frac{1}{2} E_1 E_2 E_3 \right)^{\frac{1}{3}} \qquad (5.b)$$

where

$$k_n = 1 \pm \frac{1}{n}$$

Highlight that relation of type (5.a,b) is valid for $n = 1,2,3$ only.

So, we come to the conclusion that the harmonics with $n = 1,2,3$ play a special role in energy evolution for certain.

Further reasoning will be applied to an arbitrary evolving cell (*EC*) which is assumed to be the simplest possible open system capable to evolve under the above conditions[18][19].

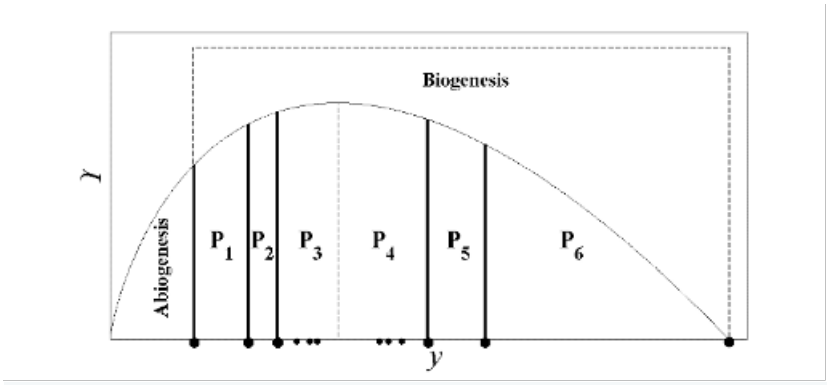

Fig. 1. *Discrete spectrum of Υ in energy evolution.*

In the plot, by an abscissa axis, an energy exchange rate*y* is indicated while by the ordinate axis an average efficiency of energy exchange Υ. Nodes Υ_n are shown by vertical segments (rightmost harmonic Υ is zero). The *PoB* is marked by the bigger black dots while the points with degenerate bifurcation by the smaller ones. In total, it composes *6* phases *Pi*. *S*tages of biogenesis and abiogenesis are also shown.

2.2. Energyaffinity in pair phase – chemical

The declared phase structure of energy evolution assumes that the energy conditions in each phase are unlikely.Then, while in different phases, *EC* experiences dissimilar energy impact. Hence, to be effective in each phase, *EC* has to adapt to altered energy conditions [20] and reorganize its energy exchange pattern (*EEPi*). Then, due to the known effect of thermodynamic (energy) affinity[21], each particular phase should obtain dissimilar support in the form of the used chemical compound, call it *Zi*.

So, the residence of *EC* in different phases should be accompanied by the activation of the distinct chemical support through the link *Pi* ↔ *EEPi* ↔ *Zi*.

2.3. Coupling of mechanisms for cell energy metabolism

Look at the process of energy evolution under the angle of mechanisms for cell energy metabolism (*MCEM*) in terms of energy efficiency and concisely highlight the well-established facts.

It is known that all cells use adenosine 5′-triphosphate (*ATP*) as the source of metabolic energy to support their energy activity. In time, *EC* developed three major *ATP*-related *MCEM* which are glycolysis (*GL*), photosynthesis (*PH*), and oxidative metabolism (*OM*)[22] that altogether work across the energy stage of biogenesis as shown in Fig. 2.

These mechanisms are different in terms of efficiency for the bidirectional energy exchange. It is known that the net energy gain of *GL* is equivalent to two molecules of *ATP* while *OM* is capable to yield up to 15-16-fold improvement of that. The energy efficiency of *PH* comes from its ubiquity and independence from consumption of the existing organics which accounting for its oxygenic form can put *PH* in an intermediate position between *GL* and *OM* [22][23][24][25].

On the other hand, all above *MCEM* manifest a high level of evolutionary coexistence. It is understood that though *GL* as anaerobic respiration arose very early in evolution, it is still quite compatible and consistent with the *PH* and *OM*, and all present-day cells are to support *GL*. Release of oxygen as a consequence of *PH*, at the same time, is a precondition for the development of *OM*. At this, some researchers believe that *OM* may have evolved before *PH* due to the activity of ancient microbes, which means oxygen was available for living entities as early as epoch of *GL* dominance[26][27][28]. Hence, to support the energy needs of *EC*, all three *MCEM* work in accordance and could replace each other if required, which indicates that the energy boundaries between *MCEM* can be considered as floating and sufficiently conditional. Also, pay attention to the actual bidirectional (inward and outward *EC*) character of *MCEM*, which is a reliably ascertained fact[29][30][31].

Summarizing above, the major *MCEM* is different in its nature, yet can work altogether across the entire energy stage of biogenesis as schematically shown in Fig. 2.

2.4. CEL and *MCEM* as two faces of energy evolution

Here, we would like to highlight that the process of energy evolution in *EC* can be described from two equivalent angles. Firstly, based on the *CEL* approach, energy evolution is the *4*-phase structure with clear boundaries, where each phase is unique in the terms of local energy conditions. Secondly, based on

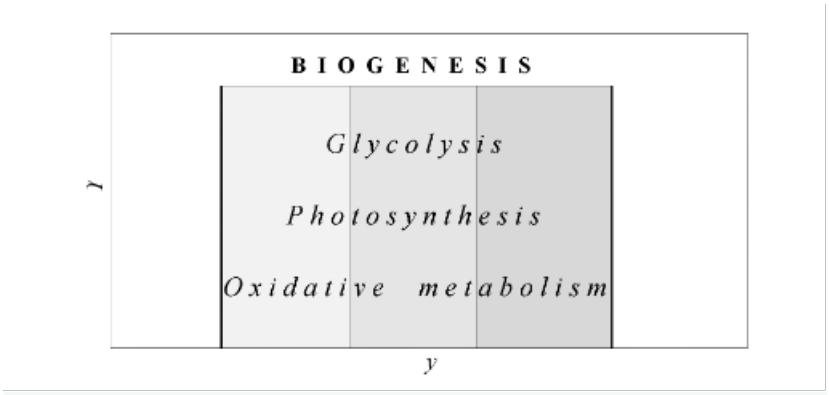

Fig. 2. *Three major MCEM are schematically shown in y - ϒ coordinates.*

In discussed approach, the exact value of energy efficiency for each mechanism as well as its *y*-positioning does not matter, so all *y*-boundaries between *MCEM* are conditional and flexible.

MCEM approach, energy evolution is the *3*-piece set of different mechanisms that are interchangeable and compatible across the entire biogenesis yet. After all, both *CEL* and *MCEM* deal with the same process.

So, we need to find the correct mathematical way to formalize the observed dual essence of the evolutionary process in *EC*.

3. Mathematical grounds of *GC* origin

3.1. Phase transformation in energy evolution

In Fig. 1, compare *y*-length of the area $P_{pos} = P1 + P2 + P3$ and $P_{neg} = P4 + P5 + P6$. So, the length of P_{neg} is $e-1$ while P_{pos} is $1 - 1/e$, which means that P_{neg} is *e* times longer than P_{pos}. Then, on average, the rate $d\Upsilon/dy$ in P_{pos} is *e* times higher compared P_{neg}. Hence, P_{neg}-phases have smeared boundaries compared to the phases in P_{pos}. It may lead to the breaking of uniqueness for the phases *P*4, *P*5, *P*6 and cause identification errors when *EC* lands in the vicinity of their border. Thus, on the one hand, we observe a clear separation between the P_{pos} and P_{neg} due to the change of the sign $d\Upsilon/dy$, on the other hand, we see vague separation within P_{neg} between the *P*4, *P*5, *P*6.

Hence, to minimize the number of identification errors, it seems reasonable that in course of evolution, the phases *P*4, *P*5, *P*6 have smeared and united into one bigger clearly identified block P_{neg}.

So, the original *6*-phase structure of energy evolution in *EC* could evolve to the more error-tolerant *4*-phase structure shown in Fig. 3.

3.2. Combinatorial mechanism of *GC* origin

Within the presented approach, the process of *EC* energy evolution reduces to, generally, an unlimited number of random leaps along the *y*-axis in *y* - Υ space. Doing this journey, *EC* can land anywhere. However, if it lands outside the range of biogenesis stage *P1* - *P6* (Fig. 1), then according to theory[15], its further evolution will be terminating.

Then, *EC* exists within a segregated energy space of *m/k* topology, in which there are *m* physically different mechanisms and *k* possible areas with dissimilar energy backgrounds where these *m* mechanisms can be randomly realized. It means that whatever changes happen, the energy status of *EC* is exclusively defined by its relative positioning with respect to the existing energy infrastructure. In other words, if *EC* does not break phase boundaries, its status does not change, so the evolutionary biography of *EC* is written relative to *m/k* energy topology.

Hence, there are a limited number of positions that *EC* can take in the *m/k* space as listed in the lower panel in Fig. 3 (b).

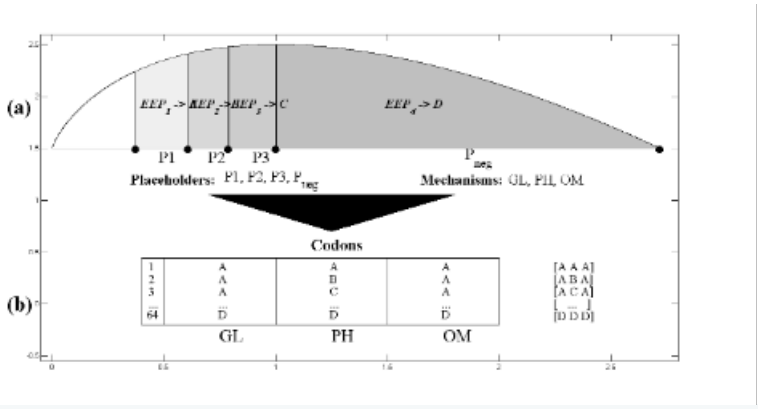

Fig. 3. *Forming of the 4-phase structure of energy evolution (a) and list of possible*

64 codon combinations (b).

In panel (a), trailing phases *P4, P5, P6* (rightmost area in dark grey) unite into one block P_{neg}.

Consequently, energy evolution can develop only if there is a coupling of k phases (nucleotides through the link $Pi \leftrightarrow EEPi \leftrightarrow Zi$) with m physical mechanisms. Saying oppositely, if there is a working intercoupling between parameters m and k of energy space, then such process is an energy evolution.

So, we come to a classic combinatorial problem on the placement of k "balls" ($P1 \leftrightarrow A$, $P2 \leftrightarrow B$, $P3 \leftrightarrow C$, $P_{neg} \leftrightarrow D$) over m "bins" (*GL*, *PH*, *OM*) shown in Fig. 4, where $m = 3$ and $k = 4$. To be exact, this is a draw for a variation with repetitions when the order of "balls" is important[32], then the total number of unique codon combinations is to be formed in accordance with

$$N = \lim_{x \to 0} V^k{}_m \qquad (6)$$

where r is the number of draw attempts, $V^k{}_m$ denotes variation with repetitions and order. At given m, k, number $N = V^k{}_m = k^m = 4^3 = 64$.

However, if the order of "balls" to pick up does not matter, which is the case for the number of building blocks (amino acids), then the total number of such amino acids is

$$N_{aa} = C'_k(m) = \left(\begin{matrix} m + k - 1 \\ k \end{matrix} \right) \qquad (7)$$

which yields $N_{aa} = 20$.

We conclude that each new evolutionary step has to meet one out of 64 possible combinations only, which split 4 available nucleotides *A, B, C, D* between the triplet sets according to the number of existing *MCEM*. Moreover, each consecutive *EC* descendent has to obey the same energy imperative that will be ultimately written into the inheritance memory of *EC* as an energy code and become mandatory for all new generations of *EC*.

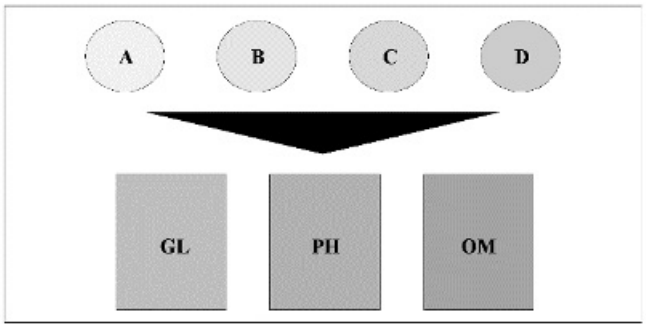

Fig. 4. *Schematic setup for draw in placement of 4 nucleotides (conditionally A, B, C, D) over 3 available mechanisms of energy metabolism in cell GL, PH, OM.*

Example. From above, record *A A A* ciphers activation of the nucleotide *A* for the consistent support firstly, of *GL,* secondly, of *PH,* and thirdly, of *OM*. In the same way, the record *A B C* signifies activation of the nucleotide *A* for the support of *GL,* then the nucleotide *B* to support *PH,* and the nucleotide *C* to support *OM*. Finally, the record *C D B* should be interpreted as original activation of nucleotide *C* at working of *GL*, then nucleotide *D* for operation with *PH*, and afterwards, activation of nucleotide *B* during operation of *OM*.

As formally the number of evolutionary attempts *r* is not limited, recording can be continued accordingly to chop nucleotides to the triplets and form the long codon chains.

4. Discussion

It is worth noting that by default, the investigated structure of *GC* belongs to the actual winners of the evolutionary competition which started a long time ago. Admittedly, all statistically significant deviations from existing *GC* were ultimately suppressed millions of years ago and we cannot say for certain what the structure of *GC* in these extinct carriers has been. That is why, to avoid speculations, we have deliberately skipped consideration of possible very original or even intermediate versions of *GC* configuration and focused on the interpretation of what we know for sure now.

So, discussion about the phenomenon of discreteness (phase separation) in genome organization across evolution is a regular process, in confirmation mention a few works[33][34][35]. Authors[33] stress that evidence for the phase separation exists across the entire tree of life while[34] underlines the principally stochastic nature of genome organization and function, the specifics of charge separation in the disordered region are discussed in [35].

The novelty of this report, predominantly, lies in the investigation of general physical (energy) and mathematical background that can support the natural transformation of a chaotic energy flow through primitive *EC* into the structured biochemical reply. To do that, an approach based on the recently confirmed phenomenon of spectral infrastructure in system energy evolution was taken. Under this angle, the triplet-based structure of *GC* could be ultimately understood as a result of intercoupling between the *CEL* and the *MCEM* approach.

In other words, we traced how the general reasoning about the configuration of energy exchange can influence to forming

of a structure of codon chain and meet the challenge of numerical constraints *3*, *4*, *20*, and *64*. We concluded that to meet the above challenge, *EC* transformation should reflect an evolutionary biography of energy changes for the winning species.

Note that the driving mechanism for *GC* structure, ultimately, is a consequence of an energy conservation law (system 1.b). On the other hand, the employed combinatorial way to formalize intercoupling between *CEL* and *MCEM* comes as a sufficiently abstract result of effective stochastic coexisting of the major procedures of energy exchange in cells.

In the information terms, we could add that due to existing laws of nature, an energy evolution to develop has to encode itself in a material carrier such as *DNA*. In the reverse way, if *GC* is fully understood and translated, it should reproduce a pattern of the sophisticated way for energy evolution in all its diversity. In light of this, it is possible to assert that biological evolution describes a natural process that transfers information from the chaotic local environment into the stable chemical known as *DNA* [36] that should be then considered primarily as the “energy code”[37].

Evidently that doing the above research, we have used some connections with the existing knowledge. In this context, indicate to the error minimization theory [38] as in this report the error minimization procedure is believed to be a driving force in the removal of unclear boundaries in trailing phases of evolution *P4*, *P5*, *P6*. Secondly, it is the stereochemical theory that is based on the phenomenon of physicochemical affinity[39][40] between anticodons and amino acids. We used these ideas to substantiate the presence of thermodynamic affinity $Pi \leftrightarrow EEPi \leftrightarrow Zi$.

In some approximation, if to ignore the obvious difference in scale, the above process of continuous to discrete transformation in energy evolution resembles the well-known development of a tree-ring segregation. Firstly, as forming of an individual pattern reflects changes in the local climate, predominantly temperature, and secondly as the possible way to retrieve these climate changes is to properly decipher existing tree-ring patterns. At this, it is worth noting that dividing of tree cells follows a highly controlled sequence of successive events described in the cell cycle. So, the emergence of the discrete solution in this example looks inescapable [41].

Pay attention that the chart in Fig. 3 (b) assumes that to cover all possible codon scenarios it is necessary to use $3 \times 64 = 192$ chart elements. At the same time, from the physical standpoint, the discussed approach does not allow an even more effective coding schema than 4^3. We mean the schemas like 8^2 with *128* elements or the exotic 64^1 with *64* elements should be rejected for the reason that the maximum number of *PoB* in our theory is limited by *6* (*3* pairs), so having *8* or *64* *PoB* just cannot fit to the existing level of understanding.

Notice that one of the advantages of the proposed concept of *GC* generation is the complete absence of any specific memory support to keep records of evolution track for *EC* as each time to create the new codon, probabilistic draw plays from scratch to place the different chemical bases in the appropriate order.

In conclusion, the above results give an understanding of the possible ways for forming the known numerical structure of *GC*. In this sense, the *64*-scenarios matrix comes as the workhorse for the practical realization of a *GC* algorithm. Nature drives this matrix over and over again, creating codon chains from the quite limited number of elements.

References

1. ^Pross, A. A. 2012. How Does Biology Emerge from Chemistry? Orig Life Evol Biosph. 42(5):433-44.
2. ^Glancy, J., Stone J.V., Wilson S.P. 2016. How self-organization can guide evolution. R. Soc. Open Sci. 3: 160553.
3. ^Luisi, P.L. 2006. The Emergence of Life: From Chemical Origins to Synthetic Biology. 1st ed. Cambridge: Cambridge University Press.
4. a, b, c, d, eMoldavanov, A. (2022) Energy Infrastructure of Evolution for System with Infinite Number of Links with Environment \\ Biosystems, 213. 104607 — doi: 10.1016/j.biosystems.2022.104607.
5. ^Sonneborn, T.M. (1965). Degeneracy of the genetic code: extent, nature, and genetic implications. In Evolving genes and proteins. ed. Bryson, V. and Vogel, H. New York: Academic Press. pp. 377–397.
6. a, bGusev, V.A., Schulze-Makuch, D. 2004. Genetic code: Lucky Chance or fundamental law of nature. Phys. Life. 1. 202- 229.
7. a, bKoonin, E.V, Novozhilov A.S. Origin and evolution of the genetic code: the universal enigma. PMID: 19117371, PMCID: PMC3293468, DOI: 10.1002/iub.146 (2009).
8. ^Crick F.H. 1968. The origin of the genetic code. J. Mol. Biol. 38:367–379. doi: 10.1016/0022-2836(68)90392-6.
9. ^Crick, F. (1990) What Mad Pursuit: A Personal View of Scientific Discovery. London: Penguin Books Ltd.
10. ^Yanofsky, C. (2007) Establishing the Triplet Nature of the Genetic Code. Cell. 128 (5): pp. 815–818. doi:10.1016/j.cell.2007.02.029.
11. ^Taylor, F.G., Coates, D. (1989) The code within codons. Bio Systems 22 (3), pp. 177-87.
12. ^Xie, P. 2017. The origin of genetic codes: from energy transformation to informatization. Biodiv. Sci. 25(1): 94-106.
13. a, b, cMoldavanov A.V. Topology of organized chaos. — 2020. — Moscow: Fizmatkniga. — p. 66. — ISBN 978-5-89155-338-5.
14. a, bMoldavanov, A. (2021) Abiogenesis-Biogenesis Transition in Evolutionary Cybernetic System \\ Cybern Syst. Advance online publication.
15. a, b, cA. Moldavanov. “Randomized Continuity Equation Model of Energy Transport in Open System,” in AMSM 2017: Proceedings of the 2017 2nd International Conference on Applied Mathematics, Simulation and Modelling, 258 – 265, 2017.
16. ^Moldavanov A.V., Analytical and Numerical Model for Evolution of Minimal Cell with Infinite Number of Energy Links. \\ In Proceedings “Mathematical biology and bioinformatics”. — 2020. — Vol 8. — Article # e16. doi: 10.17537 /icmbb20.15.
17. ^R. Swenson, Syst. Res. Behav. Sci. 6(3), 187–197 (1989).
18. ^Luisi, P. L. 2014. The Minimal Autopoietic Unit. Orig Life Evol Biosph. 44(4):335-38.
19. ^Smith J.M. The Problems of Biology (The Problems of Science). Oxford University Press: Oxford. 1986. pp. 144 pp.
20. ^Gorban, A. N., Tyukina, T. A., Smirnova, E. V., Pokidysheva. 2016. Evolution of adaptation mechanisms: adaptation energy, stress, and oscillating death. https://arxiv.org/abs/1512.03949v2.
21. ^Koukkari, P., Pajarre, R., and Kangas, P. (2018). Thermodynamic affinity in constrained free-energy systems. Monatshefte für Chemie, 149(2), 381–394.

22. [a, b]*Cooper GM. The Cell: A Molecular Approach. 2nd edition. Sunderland (MA): Sinauer Associates; 2000. The Origin and Evolution of Cells. Available from: https://www.ncbi.nlm.nih.gov/books/NBK9841/.*
23. [^]*Biology LibreTexts First Cells liensed under CK-12. Last updated 5 Mar 2021. https://bio.libretexts.org/bookshelves/introductory_and_general_biology/book%3a_introductory_biology_(ck-12) / 05%3a_evolution/5.06%3 a first_cells.*
24. [^]*Blankenship RE. Early evolution of photosynthesis. Plant Physiol. 2010 Oct;154(2):434-8. doi: 10.1104/pp.110.161687. PMID: 20921158; PMCID: PMC2949000.*
25. [^]*Bloom & Lancaster High-Efficiency Photorespiration 1 Title: "Manganese binding to Rubisco could drive a photorespiratory pathway that increases the energy efficiency of photosynthesis." Bloom & Lancaster High-Efficiency Photorespiration 1 Title Authors' Names and Affiliations: Arnold J. Bloom, Department of Plant Sciences, University of California at Davis, Davis, CA 95616; Ky.*
26. [^]*Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. Energy conversion: Mitochondria and chloroplasts. In: Gibbs S, editor. The molecular biology of the cell. 4th edition. New York: Garland Science; 2002. p. 769.*
27. [^]*https://www.sciencedaily.com/releases/2018/03/180306093304.htm.*
28. [^]*https://zeenews.india.com/science/ancient-microbes- produced-oxygen-earlier-than-thought-2087791.html.*
29. [^]*de Meis L, Grieco MA, Galina A. Reversal of oxidative phosphorylation in submitochondrial particles using glucose 6-phosphate and hexokinase as an ATP regenerating system. FEBS Lett. 1992 Aug 17;308(2):197-201. doi: 10.1016/0014-5793(92)81273-o. PMID: 1499730.*
30. [^]*doi: https://doi.org/10.1101/334458. Insights into oxygenic photosynthesis from Aurora 4 Christen L. Grettenberger1*, Dawn Y. Sumner1 , Kate Wall1 , C. Titus Brown2,3, Jonathan Eisen2 5 , Tyler J. Mackey1,4 , Ian Hawes5 , Anne D. Jungblut.*
31. [^]*On the origins of oxygenic photosynthesis and aerobic respiration in Cyanobacteria. Soo, Rochelle M, Hemp, James, Parks, Donovan H, Fischer, Woodward W., Hugenholtz, Philip. Science (American Association for the Advancement of Science), 2017, Vol.355 (6332), p.1436-1440.*
32. [^]*Ryser, Herbert John (1963), Combinatorial Mathematics, The Carus Mathematical Monographs 14, Mathematical Association of America.*
33. [a, b]*Feric, M., Misteli, T. Trends in cell biology. Phase separation in genome organization across evolution. Vol. 31, Iss. 8, 671-685. 2021.*
34. [a, b]*Sood, V., Misteli, T. The stochastic nature of genome organization and function, Current Opinion in Genetics & Development, Volume 72, 2022, Pages 45-52.*
35. [a, b]*Wang, Y., Huabin, Zhou, X. S., Huang, Q., Li, S., Zhirong, L. L., Luhua, C., Lai Z. Charge Segregation in the Intrinsically Disordered Region Governs VRN1 and DNA Liquid-like Phase Separation Robustness, Journal of Molecular Biology, Volume 433, Issue 22, 2021, 167269.*
36. [^]*Freeland, S. 2013. The Evolutionary Origins of Genetic Information. BioLogos. Online publication.*
37. [^]*Völker, J. 2020. Genetic code evolution and Darwin's evolution theory should consider DNA an 'Energy Code'. Rutgers. Online publication.*
38. [^]*Novozhilov, A.S., Koonin, E.V. Exceptional error minimization in putative primordial genetic codes. Biol Direct 4, 44*

(2009). https://doi.org/10.1186/1745-6150-4-44.

39. ^*Woese, C.R., Hinegardner, R.T. Engelberg, J. Universality in the Genetic Code. Science. 1964, 144: 1030-1031. 10.1126/science.144.3621.1030.*
40. ^*Hinegardner, R.T., Engelberg, J. Rationale for a Universal Genetic Code. Science. 1963;142:1083–1055.*
41. ^*Ursache, R., Nieminen, K., Helariutta, Y. (2013). Genetic and hormonal regulation of cambial development. Physiol. Plant. 147 36–45. 10.1111/j.1399-3054.2012.01627.x.*