P r o c e e d i n g s

of the

2nd International Conference "Theoretical Approaches to BioInformation Systems" (TABIS.2013)

September 17 – 22, 2013, Belgrade, Serbia

Editors

B. Dragovich, R. Panajotović, D. Timotijević

Golden and Harmonic Mean in the Genetic Code

Miloje Rakočević [a]
Department of Chemistry, Faculty of Science, University of Niš, Serbia

Abstract

In previous two works [1, 2] we have shown the determination of the genetic code by golden and harmonic mean within standard Genetic Code Table (GCT), i.e. nucleotide triplet table, whereas in this paper we show the same determination through a specific connection between two tables - of nucleotide doublets Table (DT) and triplets Table (TT), over polarity of amino acids.

1 Introduction

In a previous work we have shown that golden mean is a characteristic determinant of the genetic code (GC), regarding on the codons binary tree, 0-63 [1]. In a second one we showed a splitting of Genetic Code Table (GCT) into three equal and significant parts, using the harmonic mean ($H(a, b) = 2ab/(a+b); a = 63, b = 31.5$) [2]. In this paper, however, we will show that a specific unity of golden mean and harmonic mean appears to be the determinant of Rumer's Table of 16 nucleotide doublets [3] (Tables 1 & 2 in relation to Tables 3 and 4).

As we have shown, golden mean "falls" between the 38th and 39th codon (38. CAA, 39. CAG), which code for glutamine (Q), a more complex of

---

[a] e-mail address: milemirkov@open.telekom.rs

01. G	GG (6)	02. F	UU (4)	03. L		01. G	GG (6)	02. F	UU (4)	03. L
04. P	CC (6)	05. N	AA (4)	06. K		04. P	CC (6)	05. N	AA (4)	06. K
07. R	CG (6)	08. I	AU (4)	09. M		07. A	GC (6)	08. Y	UA (4)	09. St.
10. A	GC (6)	11. Y	UA (4)	12. St.		10. R	CG (6)	11. I	AU (4)	12. M
13. T	AC (5)	14. H	CA (5)	15. Q		13. V	GU (5)	14. C	UG (5)	15. W
16. V	GU (5)	17. C	UG (5)	18. W		16. T	AC (5)	17. H	CA (5)	18. Q
19. S	UC (5)	20. D	GA (5)	21. E		19. L	CU (5)	20. S	AG (5)	21. R
22. L	CU (5)	23. S	AG (5)	24. R		22. S	UC (5)	23. D	GA (5)	24. E

Table 1. Rumer's Table of nucleotide doublets   **Table 2.** The modified Rumer's Table

~~Table 1~~: Distributions of AAs after nucleotide doublets presented in Table 2: Four squares with dark tones (outer) contain four first doublets from Table 2 and four light (inner) contain four second doublets. In amino acids (within their side chains) at outer/inner areas there are 369/369 nucleons and 61/61 atoms, respectively. All AAs in outer area are nonpolar whereas those in inner area are polar, measured by cloister energy.

only two amide amino acids (AAs); two codons, adjacent to the codons (40.UGU, 41.UGC), which code for one of the only two sulfur AAs, cysteine (C). This "harmonization" of diversity is increased by the harmonic mean, in position 42 on the sequence 0-63. The harmonization extends further to "stop" codon (42.UGA) and to codon (43.UGG) that codes for the most complex AA, tryptophan (W). (The "42" as ending position on the "Golden route" - with Fibonacci numbers - on the Farey tree, corresponding with six-bit GC binary tree [1].)

On the other side, the splitting of GCT into three parts through harmonic mean [2] makes that AAs are distinguished on the basis of the validity of the evident regularities of key factors, such as polarity, hydrophobicity and enzyme-mediated AAs classification (with parameter values as in Table 2.1 in Rakočević, 2013).

2   A new rearrangement of nucleotide doublet Table

With a minimal modification of Rumer's nucleotide doublets Table (DT) follows the next result: if at the beginning of first sub-system[1], with 6/4 hydrogen bonds, are GG/UU doublets, chemical reasons require GU/UG doublets at the beginning of the second sub-system, with 5/5 hydrogen bonds, instead of AC/CA as it is in Table 1. From the same reasons, we have the changes: CG/GC & UC/CU on the left and AU/UA & GA/AG on the right. With the four first doublets we have four outer squares, i.e. codon families (n_1 = GG,UU,GU,UG) in nucleotide triplets Table (TT) which code for nonpolar AAs; the four second doublets give four inner codon families (n_2 = CC,AA,AC,CA), which code for polar AAs (Table 3)[2]. With the four third (n_3 = GC,CU;UA,AG) and four fourth (n_4 = CG,UC;AU,GA) doublets are chosen eight intermediate codon families, the first three code for nonpolar AAs [(CU,AU,GC) → (L, I, M, A)] and the second five for polar AAs [(UC,UA,CG, AG, GA) → (S,Y, R, D, E)] (Table 4). As we see the polar/nonpolar distribution of squares in Table 3 is realized as 4±0 and in Table 4 as 4±1. By this, the polarity/nonpolarity is taken after cloister energy as in Ref. [4][3].

3   Particles number balances through polarity

Distinctions through polarity, presented in Tables 3 & 4, are followed by the balance of the number of nucleons and atoms. Irrespectively of the Table of nucleotide doublets, V. shCherbak showed [Ref. 10, Fig. 10, p. 173] that it makes sense to display the Table of nucleotide triplets (TT) exactly as here in Tables 3 & 4: four squares at the corners and four squares in the center as in Table 3; then, eight squares in middle, i.e. in "between" positions, as in Table 4. After shCherbak's view, the balance of the number of nucleons is the next: AAs in four squares in the corners of TT as well as AAs in four

---

[1]Two subsystems, each with two quadruplets; in total four quadruplets: two on the left as one-meaning (each nucleotide doublet codes for one AA) and two on the right as two-meaning (each nucleotide doublet codes for two AAs or for one AA and termination in the protein synthesis).

[2]Histidine (H) is neutral in cloister energy with the value ± 0 [4], but polar in hydropathy [5], polar requirement [6, 7] and in hydrophobicity [8, 9].

[3]Cloister energy is "a formal free energy (= cloister energy) of transfer of the amino acid from the outside of a protein to the inside   I use cloister energy in preference to other measures of amino acid hydrophobicity-philicity because it is an in situ measure of the property of interest" [4].

1st	2nd letter				3rd
	U	C	A	G	
U	UUU UUC **F** UUA UUG **L**	UCU UCC UCA **S** UCG	UAU UAC **Y** UAA UAG **CT**	UGU UGC **C** UGA **CT** UGG **W**	*U* *C* *A* *G*
C	CUU CUC CUA **L** CUG	CCU CCC CCA **P** CCG	CAU CAC **H** CAA CAG **Q**	CGU CGC CGA **R** CGG	*U* *C* *A* *G*
A	AUU AUC **I** AUA **M** AUG	ACU ACC **T** ACA ACG	AAU AAC **N** AAA AAG **K**	AGU AGC **S** AGA AGG **R**	*U* *C* *A* *G*
G	GUU GUC **V** GUA GUG	GCU GCC **A** GCA GCG	GAU GAC **D** GAA GAG **E**	GGU GGC **G** GGA GGG	*U* *C* *A* *G*

1st	2nd letter				3rd
	U	C	A	G	
U	UUU UUC **F** UUA UUG **L**	UCU UCC UCA **S** UCG	UAU UAC **Y** UAA UAG CT	UGU UGC **C** UGA CT UGG **W**	U C A G
C	CUU CUC CUA **L** CUG	CCU CCC CCA **P** CCG	CAU CAC **H** CAA CAG **Q**	CGU CGC CGA **R** CGG	U C A G
A	AUU **I** AUC AUA AUG **M**	ACU ACC ACA **T** ACG	AAU AAC **N** AAA AAG **K**	AGU AGC **S** AGA AGG **R**	U C A G
G	GUU GUC GUA **V** GUG	GCU GCC GCA **A** GCG	GAU GAC **D** GAA GAG **E**	GGU GGC GGA **G** GGG	U C A G

Table 3: Distributions of AAs after nucleotide doublets presented in Table 2: Four squares with dark tones (outer) contain four first doublets from Table 2 and four light (inner) contain four second doublets. In amino acids (within their side chains) at outer/inner areas there are 369/369 nucleons and 61/61 atoms, respectively. All AAs in outer area are nonpolar whereas those in inner area are polar, measured by cloister energy.

squares in the center of TT have 369 nucleons [(F91 + L57 + V43 + G01 + W130 + C47 = 369); (P41 + T45 + K72 + N58 + Q72 + H81 = 369)].

To this shCherbak's insight, we now add: the same quantity give the AAs in right site of TT; the right site in relation to the diagonal F-G[4] in TT (S31+Y107+R100+S31+R100 = 369). On the left side of the diagonal there are 336 nucleons (L57+I57+M75+A15+D59+E73=336), what means 33 nucleons less, in relation to 369. With this emergence of difference of "33" on the scene appears a specific self-similarity because the number 33 is an important determinant of the number of atoms in the rows and columns of GCT, i.e. of TT[5].

To this self-similarity determination by nucleon number we now also add the self-similarity determination by atom number: AAs in four squares in the corners as in the center of TT have 61 atoms in AA side chains [(F14+L13+V10+G01+W18+C05=61); (P08+T08+K15+N08+Q11+H11=61)]. In relation to the diagonal F-G, in TT, there are 58 and 59 atoms, respectively; on the left: L13+I13+M11+A04+D07+E10=58, and on the right: S05+Y15+R17+S05+R17=59. These quantities (58 and 59) are the same as the quantities of hydrogen atoms in Sukhodolets' system (what is a further self-similarity): 58 in two inner and 59 hydrogen atoms in two outer rows [Ref. 11, Tab. 7, p. 830], [12]. (Notice that 58 + 59 = 117 is total number of hydrogen atoms in 20 canonical AAs of GC, within their side chains, what is the self-similarity once more.)

4   Determination through "golden whole"

The splitting of GCT (i.e. TT) into 4 outer, 4 inner and 8 intermediate squares, corresponding to responsible nucleotide doublets, leads us to the following conclusion. Within the set of all n-gons, where n is even number, the case n = 4 is only and one case where harmonic mean of "golden whole" $(n^2 - n)^6$ and its

---

[4]Starting from diagonal (F-G) together with two adjacent ones (S-R and L-E) we give a central space ("in" space) in a strict balance with the set of polar AAs, polar through hydropathy [5]: atom number in two and two sets ("in", "out" / polar, nonpolar) differs exactly for ±1 and ±10, respectively [Ref. 11, Tab. A.3, p. 840 in relation to equations 11-14, p. 838].

[5]If we consider the set of "61" of AAs, then in two rows, YNR & RNY, there are 8 x 33 and in two other, YNY & RNR, 10 x 33 atoms. On the other hand, in two pyrimidine columns, NYN, there are (9 x 33) - 1 and in two purine ones, NRN, (9 x 33) + 1 of atoms [Ref.12, Tab. 3a, p. 224]. (Y for pyrimidine, R for purine and N for all four types of nucleotides.)

[6]The full equation for "golden whole" can have the next form: $n^2 \pm n$ = "WHOLE"; for n = 4, the difference is 12 and the sum 20. In relation to genetic code there are self-similarities expressed through mathematical operations: (4 x 4) - 4 = 12; (4 x 4) + 4 = 20, (4 x 4) x 4 = 64. [Number 12, correspondent to 4+4+4 doublets in sequence $[n_1 - (n_3 \, or \, n_4) - n_2]$; number 20 as 20 AAs (4 x 4 = 16 AAs of alanine stereochemical

half $[(n_2 - n)/2]$ equals $2n$, and $n^2 - n = 3n$. So, in this case we have that the ratio 2:3 appears to be the harmonic mean within the harmonic mean, and, by this, the sequence $n_1 - (n_3 \, or \, n_4) - n_2$ corresponds with the Cantorian triadic set. Moreover, such a harmonic mean appears to be corresponding with the number of "small squares" within intermediate space in form of only one "ring" as it follows: $[(2 + 2) + (4 \times 0) = 4]$; $[(4 + 4) + (8 \times 1) = 16]$; $[(6+6) + (12 \times 2) = 36]$; $[(8+8) + (16 \times 3) = 64]$ etc. As it is self-evident, the symmetrical "out - middle- in" arrangement (1:1:1 of rings) is not possible for $n \neq 4$, neither for n-gons nor for n-letter alphabets. At the same time here is a self-similarity expressed through the number of "small squares" in the sequence $n_1 - (n_3 \, or \, n_4) - n_2$ and the number of codons within them: four squares per n_1, n_2, n_3, n_4, each square per four codons. Moreover, there is a self-similarity between golden and harmonic mean versus 4-letter alphabet: 1n as 1 square (1 nucleotide doublet), 2n as harmonic mean (in the sense above said), $3n$ as golden whole and $4n$ as the sum $n_1 + n_2 + n_3 + n_4$; all these versus 1 letter of alphabet (as letter minimum), 2 letters as word root (nucleotide doublet), 3 letters as 3-letter word (codon) and 4 letters as letter maximum within alphabet[7].

5   Concluding remarks

With the title of the paper is given a working hypothesis that the golden mean (GM) and harmonic mean (HM) are determinants of the genetic code. The findings presented by four illustrations show that this hypothesis is confirmed. However, unlike the previous access to the same determination, it refers not only the analysis of the nucleotide triplet Table, but rather refers to the two tables - Table of doublets (DT) and Table of triplets (TT). In fact, it is precisely presented that these tables are unique in terms of determination just over GM and HM. It is expected that all these uniqueness correspond to the same, or similar uniqueness, found by other authors ([16, 17, 18, 19]), what in future researches should be checked. However, presented facts are such that ones reaffirm the other and vice versa. All together, they favor the recognition that the chemical reactions that determine the GC are not only the reactions in a "test tube", but these reactions are associated with a specific balance of the number of particles (atoms and nucleons); balance, determinated by unique arithmetic and algebraic regularities and expressed in the form of specific (nonfractal) self-similarity ("a harmonized chemistry"). From this it follows further that presented facts also support the hypothesis that the genetic code was from very begining, in prebiotic times and conditions, a complete code [10, 15]. On the other hand, the knowledge that "the chemistry of living" is actually a harmonized chemistry requires great care in medicine, agriculture and natural environment, taking into account the fact that this harmonization is

---

types and 4 of non-alanine stereochemical type); number 64 as 64 codons.] (About four stereochemical types see Ref. [13] and [14].)

[7]By this one must notice that all these self-similarities are possible only for 4-letter alphabet and 3-letter words.

strictly immanent to the living as such, mediated by genetic code as such.

Acknowledgments

References

[1] M. M. Rakočević, "The genetic code as a Golden mean determined system," *Biosystems* **46**, 283–291 (1998).

[2] M. M. Rakočević, "Harmonic mean as a determinant of the genetic code," arXiv:1305.5103v4 [q-bio.OT] (1998).

[3] Yu. B. Rumer, "O sistematizacii kodonov v geneticheskom kode," *Dokl. Akad. Nauk. SSSR* **167**, 1393–1394 (1966).

[4] R. Swanson, "A unifying concept for the amino acid code," *Bull. Math. Biol.* **46**, 187–207 (1984).

[5] J. Kyte and R.F. Doolittle, "A simple method for displaying the hydropathic character of a protein," *J. Mol. Biol.* **157**, 105–132 (1982).

[6] C.R. Woese et al., "On the fundamental nature and evolution of the genetic code," in *Cold Spring Harbor Symp. Quant. Biol.* **31**, 723–736 (1966).

[7] B. G. Konopel'chenko and Yu. B. Rumer, "Klassifikaciya kodonov v geneticheskom kode," *Dokl. Akad. Nauk. SSSR* **223**, 471–474 (1975).

[8] S. D Black and D. R Mould, "Development of hydrophobicity parameters to analyze proteins, which bear post- or cotranslational modifications," *Anal. Biochemistry* **193**, 72–82 (1991).

[9] V.R. Chechetkin and V.V. Lobzin, "Stability of the genetic code and optimal parameters of amino acids," *J. Theor. Biol.* **269**, 57–63 (2011).

[10] V. I. Shcherbak, "The arithmetical origin of the genetic code," in: *The Codes of Life* (Springer, 2008).

[11] M. M. Rakočević, "Genetic code as a coherent system," *Neuroquantology* **9** (4), 821–841(2011); (www.rakocevcode.rs)

[12] V. V. Sukhodolets, "A sense of the genetic code: reconstruction of the prebiological evolution stage," *Genetics* **XXI** (10), 1589–1599 (1985) [in Russian].

[13] E. M. Popov, *Strukturnaya Organizaciya Belkov* (Nauka, Moscow, 1989) [in Russian].

[14] M. M. Rakočević and A. Jokić, "Four stereo chemical types of protein amino acids: synchronic determination with chemical characteristics, atom and nucleon number," *J. Theor. Biol.* **183**, 345–349 (1996).

[15] M. M. Rakočević, "A harmonic structure of the genetic code," *J. Theor Biol.* **229**, 463–465 (2004).

[16] B. Dragovich, "*p*-Adic structure of the genetic code," *NeuroQuantology* **9** (4), 716–727 (2011); arXiv:1202.2353v1 [q-bio.OT].

[17] T. N gadi, The Multiplet structure of the genetic code, from one and small number, *Neuroquantology* **9** (4), 767–771 (2011).

[18] N. Ž. Mišić, "Nested numeric/geometric/arithmetic properties of shCherbak's prime quantum 037 as a base of (biological) coding/computing," *Neuroquantology* **9** (4), 702–715 (2011).

[19] F. Castro-Chavez, "The quantum workings of the rotating 64-grid genetic code," *Neuroquantology* **9** (4), 728–746 (2011).

Note added in Proof:

Table 3 is on p. 208 and Table 4 on p. 209.

Legend to Table 3:

Table 3. Distributions of AAs after nucleotide doublets presented in Table 2: Four squares with dark tones (outer) contain four first doublets from Table 2 and four light (inner) contain four second doublets. In amino acids (within their side chains) at outer/inner areas there are 369/369 nucleons and 61/61 atoms, respectively. All AAs in outer area are nonpolar whereas those in inner area are polar, measured by cloister energy.

Legend to Table 4:

Table 4. Distributions of AAs after nucleotide doublets presented in Table 2: Four squares with dark tones contain four third doublets from Table 2 and four light contain four fourth doublets; two and two doublets on the right, and two and two on the left. In amino acids (within their side chains), in right/left areas there are 369/369-33 nucleons and 59/58 atoms, respectively. All AAs on the right together with D & E from the left are polar and other on the left are nonpolar. Because the balance is realized in relation to diagonal and not in relation to the type of nucleotide doublets (third or fourth) it follows that positions 3 & 4 in DT are not invariant, but only positions 1 & 2 as it is shown in Table 3.